REPORT DOCUMENTATION PAGE

AFRL-SR-BL-TR-98-
0375

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	1995	Final

4. TITLE AND SUBTITLE
Biochemical and Genetic Characterization of the Transcription Elongation Factor TFIIS from the Yeast Saccharomyces cerevisiae

5. FUNDING NUMBERS

6. AUTHORS
Karen Renee Christie

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of California, Berkeley

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NI
110 Duncan Avenue, Room B-115
Bolling Air Force Base, DC 20332-8080

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
See attached.

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

Biochemical and Genetic Characterization of the Transcription Elongation Factor TFIIS from the Yeast *Saccharomyces cerevisiae*

by

Karen Renee Christie

B.A. (University of Tennessee, Chattanooga) 1989
B.S. (University of Tennessee, Chattanooga) 1989

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Genetics

in the

GRADUATE DIVISION

of the

UNIVERSITY of CALIFORNIA, BERKELEY

Committee in charge:

Professor Caroline M. Kane, Chair
Professor Jasper D. Rine
Professor Hillary C. M. Nelson

1995

The dissertation of Karen Renee Christie is approved:

Caroline M. Kane 12-4-95
Chair Date

12-14-95
Date

12-15-95
Date

University of California, Berkeley

1995

ABSTRACT

Biochemical and Genetic Characterization of the Transcription Elongation Factor TFIIS from the Yeast *Saccharomyces cerevisiae*

by

Karen Renee Christie

Doctor of Philosophy in Genetics

University of California, Berkeley

Professor Caroline M. Kane, Chair

Regulation of the process of transcriptional elongation is an important control mechanism in the expression of some genes. To fully understand this form of regulation will require better understanding of the functions of transcription elongation factors. The goal of this work was to characterize the transcription elongation factor TFIIS from *Saccharomyces cerevisiae,* originally called P37. I demonstrated that, like the mammalian TFIIS proteins, the yeast protein stimulates RNA polymerase II to cleave the nascent RNA transcript and to read-through an intrinsic block to elongation. Investigation of the protein:protein contacts between TFIIS and RNA polymerase II indicated that the carboxyl-terminal domain of the largest subunit, subunit four, and subunit seven of the polymerase are not required for TFIIS to promote cleavage and read-through by the polymerase. In addition, the carboxyl-terminal half of the yeast TFIIS protein is sufficient for both of these *in vitro* activities. This result is consistent with the previous results demonstrating the carboxyl-terminus of mouse TFIIS was sufficient to activate RNA polymerase *in vitro.*

Very little is known about the function of TFIIS *in vivo*. Cells lacking this activity are viable but sensitive to the drug 6-azauracil, which causes drops in the cellular levels of both GTP and UTP. I demonstrated that TFIIS lacking the amino terminus is sufficient to rescue this phenotype. To identify proteins that interact with TFIIS *in vivo*, I chose the two-hybrid approach. A preliminary screen suggested that this approach is feasible despite the interaction between TFIIS and RNA polymerase II *in vitro*. This screen identified two proteins as putatively interacting with TFIIS. One protein was *MSN4*, a zinc-finger protein involved in carbon utilization. The second had homology to the mouse proteins Smcy and Smcx and to the human proteins SMCX and retinoblastoma binding protein 2. Whether these proteins actually interact with TFIIS remains to be demonstrated.

Caroline M. Kane

To my great aunt, Lillian Welton

and

to my grandmother, Jessie Donaldson

ACKNOWLEDGMENTS

In some ways these are probably the most important pages in this thesis, because I could not have done this alone. There are many people who have helped my along the way and here, and I would like to give some of the credit where it is due. There are many other people whom I have interacted with along the way who deserve thanks, but whom I have not included due to lack of space or memory. Many thanks.

I thank Caroline Kane for getting me started. When I begin in the lab, she was a great resource to me, both technically and intellectually. She has also been a good friend and advisor, scientifically and otherwise.

Many thanks to my thesis committee. They have been everything I could have asked for in a thesis committee and more. Jasper Rine, who introduced me to Caroline, advised me to use studying for my prelim as an opportunity to learn about a few areas that interested me and to enjoy the process. It may have been the single most useful comment made to me during graduate school and proved equally useful in writing this thesis. Thanks also for a timely reminder on the elements of style. Many thanks to Hillary Nelson, who has been an inspiration and who encouraged my aspiration to do a post-doc in a foreign country. Over the years, they have been interested, supportive, and critical. I thank them both for giving me their time and attention.

I thank my father for always being there for me.

I thank my mother, who was always willing to answer my questions, even when I was five and wanted to know about atoms and molecules and DNA.

I cannot adequately express my gratefulness to Mike and Alice Welton and to Susan and Alex Zandra Quinn for sharing their beautiful home with me while I wrote much of this monster. Thanks to them, I went swimming in the creek with the dogs, turtles, and fish, and the occasional otter, garter snake, or leech. I watched the egrets, the herons (blue or brown), the kingfishers, various unidentified birds, and the turkey vultures circling overhead. I became friends with Trouble the cat who followed me down to the creek one night just to sit in my lap and who sometimes sat in my lap for hours while I worked at the computer. Living off of Putah creek in the summer produced the interesting phenomenon of smashing bugs on my monitor at night and the subsequent development of the "bug-proof" (i.e. less light emitting) color scheme designed for night. I thank Mike and Alice for introducing me to the unexpected joy of hurling rotten vegetables out the back door to smash on the rocks below. Thanks to Susan's grandfather Dale, I spent the summer on the "thesis diet of champions": squash, onions and tomatoes. And I thank Susan and Alex for teaching me about grace and what really matters.

I thank Steve Mack for always caring, for always listening to me, and for always wanting to make me happy.

Many thanks to my good friend Ken Stedman for many thing, for commiserating, for drinking beer, for subscribing to San Francisco Ballet, for skiing, and just talking.

Many thanks to Mark Groudine for sharing his insight into the regulation of the c-*myc* gene. It greatly clarified the often confusing maze of literature concerning the transcription of this gene.

I feel a great appreciation for Charles Darwin. Not only did *The Origin of Species* forever change the way people think about biology, but reading it this past year brought back to me all of the reasons for why I wanted to be involved in the science of biology. It is amazingly complex and fascinating.

Now I move on to members of the lab, the place where I have spent so much of my time and energy. These are the people who have made the days worthwhile. These little blurbs are but pale shadows.

Many thanks to:
Susan Uptain, with whom I have explored the far reaches of the galaxy (and blown up a lot of imaginary rocks), toured the BioRad parking lot for a few hours, learned more than I ever imagined about the workings of the personal computer, discussed a great deal of interesting science (usually hers), beaten at only one video game (Minesweeper), and whom I have sniffed phenol for.

Zoë Weaver, who doesn't like chocolate but loves carrot cake, the only person who has made me a birthday cake since I was ten, and a very good friend.

Dan Mytelka, whom I have sat next to for the last five years, but only just learned that he is a great Magician. Without his willingness to listen to my technical problems and provide a clear, thoughtful perspective and good advice, I would never have finished Chapter 3. I will miss his insightful view and his dry humor, which pops up at the oddest moments.

Rod Weilbaecher, who possesses an amazing analytical mind, who played Badmotorfinger really loud until I had to love it, one of my best sounding boards for experimental ideas and methods of quantitation, and fellow connoisseur of beer.

Chris Seidel, who brought Magic™ to the lab, and who is leading the lab into the cyberworld (via the World Wide Web), with whom I have had deep and interesting conversations in the wee hours of the morning.

Nell Shimasaki, maker of the most amazing double chocolate brownies in existence and who reminded me of what is was like when I was just beginning my project. A most especial thanks for her motivational words during this last year.

Kristy Johnson, whom I thank for sharing with me her enthusiasm for learning. The experience of helping her to become scientifically independent in the lab is one of the most fulfilling experiences I had during graduate school.

Judy Davie, fellow member of the "yeast group" who tells a really great story about a woodpecker and who shares an appreciation for the stunningly strong mixed drinks of the South.

Linda d'Ari, who shares my enthusiasm for the art of ballet, and whose cautions on the difficulties of enzyme purification I should have listened to earlier.

Meghan Bowser and Rachel Fish, current rotation students in the lab for reminding me what it's like to be a first year, Meghan with her boundless enthusiasm and Rachel with her horror stories of the Genetics 240 final exam.

Ross Chambers, for greatly simplifying the task of finding the most appropriate references in the expanding field of research into the function of the CTD of RNA polymerase II.

Sandra Milan, "Queen of Cleavage" and the fashion queen of the lab (since Gretchen left), with whom I enjoyed the Picasso museum and people watching in Paris, and the only person with whom I can go to Long Life Veggie House regularly.

Tracy Johnson, who broadened my horizons on music (Oletta Adams and Anita Baker), says I have expanded hers as well (Pearl Jam, Sarah McLachlan, and "Freebird"), who shares an occasional desire to play the same song three times in a row (when no one else is around), and with whom I have had some stimulating conversations on cleavage and other subjects.

Dorothie Erie, a boisterous presence in the lab who shared with me some words of wisdom.

Dave Solow-Cordero, who is in part responsible for my current addiction to the sport of downhill skiing, whose taste in music I have a lot in common with, whose prelim proposals and thesis chapters I read for hours, and a good friend, despite or perhaps because of his blunt honesty.

Curtis Altmann, whom I learned about quite a bit about computers from, or sometimes in spite of, who taught me many techniques very well when he was in the mood, an opinionated man of much shouting at times, a brewer of some damn good beer, and a member of a great hiking party through the Yosemite back country.

Laura Williams, one of the most pessimistic and also one of the most analytically critical people I have ever met, who tells amazing stories about her family or about arguing with airline companies, a woman who may have missed her calling as a brilliant litigator.

Gretchen Rice, whom I looked up to when I had just joined the lab because she asked the most insightful questions and produced the prettiest RNA gels.

Dan Mirel, a fearsome practice pre-lim questioner and the source of my infection with playing the game Minesweeper.

Chris Surratt and Lakshmi SivaRaman, the calm steady presences in room 610 when I first joined the lab, but whose work contributed greatly to the feeling of excitement in the lab because it was helping to redefine the way people think about RNA polymerases.

Carrie Chen and Veronica Lustre, undergraduates who taught me how to do nearly everything when I had just joined the lab.

and many thanks to others who have passed through the lab whom I have not mentioned specifically.

Karen Renee Christie was supported in part by a National Defense Science and Engineering Graduate Fellowship.

TABLE OF CONTENTS

Dedication iii

Acknowledgments iv

Table of Contents viii

Chapter 1 - Introduction 1

The transcription cycle 2

Regulation of Transcription Elongation in Prokaryotes 4

Regulation of Transcription Elongation in Eukaryotes 7

Importance of Regulating Transcriptional Elongation in Eukaryotes . . . 8

Locations of Blocks to Elongation 14

Techniques for the identification of blocks to elongation 14

Blocks to elongation in the 3'-regions of genes 20

Blocks to elongation within genes 24

The Effects of Initiation Upon Elongation 30

Transcription Elongation Factors 35

Early Biochemistry of TFIIS proteins (mouse S-II, yeast P37) 37

Cleavage of the Nascent RNA Transcript 41

The Structure of TFIIS 44

RNA polymerase II 45

Interactions between TFIIS and RNA polymerase II 48

TABLE 50

Chapter 2 - Comparison of Full Length and Truncated Yeast TFIIS with Several Variants of RNA polymerase II *in vitro* 54

INTRODUCTION 55

MATERIALS AND METHODS 59

RESULTS 69

TFIIS Activity in *Saccharomyces cerevisiae* 69

Read-through by Purified TFIIS 70

Cleavage of Nascent Transcripts in Ternary Complexes 72

DISCUSSION 76

FIGURES 80

Chapter 3 - Further Deletion Analysis of the Amino-terminus of Yeast TFIIS, *in vitro* and *in vivo* 92

INTRODUCTION 93

MATERIALS AND METHODS 97

RESULTS 115

Purpose 115

Subcloning Strategy 116

Complementation of 6-azauracil Sensitivity *in vivo* by TFIIS Variants 117

Incomplete Complementation by the Δ2-130:P292L TFIIS Protein 120

Western Analysis of TFIIS Expression Levels 122

Summation of *in vivo* Results 123

Activities of Variant Yeast TFIIS Proteins *in vitro* 124

Stimulation of Read-through of an Intrinsic Block to Elongation 124

Alternative Quantitation of Read-through Assays 126

Stimulation of Cleavage of the Nascent Transcript 127

Sequence Analysis of the Flanking Regions of the *PPR2* gene 128

Sequence Analysis the Promoter Region of the *PPR2* gene 129

DISCUSSION 131

TABLES 136

FIGURES 142

Chapter 4 - Feasibility Study on The Use of the Two-Hybrid Technique to Identify Proteins Which Interact with Yeast TFIIS *in vivo* 172

INTRODUCTION 173

MATERIALS AND METHODS 179

RESULTS 189

Purpose 189

Construction of plasmid encoding fusion of *GAL4* DNA binding domain to TFIIS 190

Detection of Expression of the Fusion Protein from pKC14a-1 190

Reporter Strains 191

Assays for Activation of the Three Integrated Reporter Constructs 192

Comparison of Reporter Activation by DBD-TFIIS in Y153 and Y166 194

The Effects of the Growth Media on Activation of Reporters in Y166 197

Comparison of the *URA3* and *HIS3* reporters in Y166 199

Trial Screens of YL2 Library in Y166 Expressing DBD-TFIIS . . . 201

Comparison of Efficiency calculations on SC - leu or on SC - his, leu, trp 202

Activation of Reporter Genes in Trial Screens 202

Further Characterization of Putatively Positive pYL2 Plasmids 205

DISCUSSION 208

TABLES 213

FIGURES 217

REFERENCES 245

LIST OF TABLES

Chapter 1

Table 1 - Genes Reported to Contain Blocks to Transcriptional Elongation 50

Chapter 3

Table 1 - List of Strains 136

Table 2 - Sequencing Primers for Region Containing *PPR2* 138

Table 3 - Doubling Times of KJY1 Strains with or without 6-azauracil 140

Chapter 4

Table 1 - List of Strains 213

Table 2 - Results of Preliminary Activation Domain Library Screens 215

LIST OF FIGURES

Chapter 2

Figure 1 - Protein preparations used in transcription assays 80

Figure 2 - Sequence of the Tailed Template pCpGEMTerm 82

Figure 3 - Time Course of TFIIS-Stimulated Read-through of Intrinsic Blocks to Elongation 84

Figure 4 - Comparison of Ability of the Full Length and Truncated TFIIS Proteins to Stimulate Read-through by RNA polymerase II 86

Figure 5 - Quantitation of Read-through Activity by Variant Forms of RNA polymerase II 88

Figure 6 - TFIIS-Stimulated Cleavage of Nascent RNA in Ternary Complexes 90

Chapter 3

Figure 1 - Polylinker of pKC-RS314 142

Figure 2 - Insert of pKC3/pKC6 and cloning strategies for pKC15 and pKC16 144

Figure 3 - Inserts and Relevant Sites of pKC15 and pKC16 146

Figure 4 - Activities of Variant Yeast TFIIS Proteins 148

Figure 5 - Complementation of 6-azauracil Sensitivity by TFIIS variants 150

Figure 6 - Growth Curves of KJY1 Strains with or without 6-azauracil 153

Figure 7 - Western to Detect TFIIS Expression in KJY1 Strains Bearing Plasmids 155

Figure 8 - Preparations of Histidine-tagged Yeast TFIIS proteins 157

Figure 9 - Read-through Stimulated by Variant Yeast TFIIS proteins . . 159

Figure 10 - Quantitation of Read-through Stimulated by Variant TFIIS proteins 161

Figure 11 - Intensities of the Run-off and TIa Bands in Read-through Assays 164

Figure 12 - Time Course of Transcript Cleavage Stimulated by Yeast TFIIS 166

Figure 13 - Promoter Regions of *RNA15* and *PPR2* 168

Figure 14 - Alignments of TFIIS Proteins from *S. cerevisiae* and HeLa 170

Chapter 4

Figure 1 - Diagram of Scheme for Two Hybrid Screen 217

Figure 2 - Diagram of pKC14a-1 219

Figure 3 - Comparison of *HIS3* and *URA3* reporters of Y166 with *HIS3* Reporter of Y153 221

Figure 4 - Comparison of LacZ reporters in Y153 and Y166 223

Figure 5 - Growth of Y166 + pKC14a-1, pNI12 and Y166 + pAS-SNF1, pNI12 on Various Selective Media 225

Figure 6 - Assays for β–Galactosidase Activity of Y166 + pKC14a-1, pNI12 and Y166 + pAS-SNF1, pNI12 on Various Selective Media 227

Figure 7 - Growth of Y166 + pKC14a-1, pNI12 or Y166 + pKC14a-1, pNI12 and Y166 + pKC14a-1, pNI12 on Various Selective Media 229

Figure 8 - LacZ assays of Y166 + pKC14a-1, pNI12 or Y166 + pKC14a-1, pNI12 and Y166 + pKC14a-1, pNI12 on Various Selective media 231

Figure 9 - Y166 + pKC14a-1 Transformed with YL2 Library DNA 10 Days Growth on SC - his, leu, trp, ura 233

Figure 10 - β–galactosidase Assays of Transformations and Control Strains 235

Figure 11 - Growth of Putatively Positive pYL2 Plasmids 237

Figure 12 - β-Galactosidase Assays of Cells Containing Putatively Positive pYL2 Plasmids 239

Figure 13 - Restriction Analysis of Putatively Positive pYL2 Plasmids 241

Figure 14 - Alignment of Predicted Amino Acid Sequence from pYL2-1B 243

Chapter 1

Introduction

All living organisms, whether multicellular forms or single cells, or viruses, contain genetic information. Genes control everything from the development of morphological structures to the complex chemistry within the cell. Genes are thus the ultimate substrate of natural selection and evolution. Understanding the regulation of gene expression is crucial for understanding life.

The expression of a gene is a complex process. In a simple view, the first step is transcription. In organisms for which DNA is the genetic material, RNA polymerase makes a messenger RNA (mRNA) transcript from a gene. A ribosome translates the mRNA transcript into a polypeptide. The resulting protein then performs some function for the cell. The protein's role in the cell may be enzymatic, structural, or regulatory. The regulation of a gene's expression may occur by modulating the level of transcription, the stability of the mRNA, the translation of the mRNA, or the stability of the protein. However, the expression of every gene is controlled, at least in part, by transcription.

The Transcription Cycle

Eukaryotic cells contain three nuclear RNA polymerases (reviewed in Sentenac, 1985; Thuriaux and Sentenac, 1992). RNA polymerase I transcribes the large ribosomal RNAs, and RNA polymerase III transcribes transfer RNAs, 5S RNA, and other small RNAs. RNA polymerase II is responsible for transcription of messenger RNAs and some small nuclear RNAs of snRNPs (reviewed in Sawadogo and Sentenac, 1990). Thus RNA polymerase II must recognize the great diversity of promoter structures that provide differential expression of many different genes in a cell. The control of transcription initiation by RNA polymerase II has been a major focus of investigations into gene regulation. The RNA polymerase must find

the right place on the chromosome to begin transcription. The timing and frequency of this event may be regulated with respect to the cell cycle, the tissue type, the stage of development, or in response to environmental signals. A great deal has been learned about the processes and proteins which control transcription initiation (reviewed in Conaway and Conaway, 1993; Tjian, 1995; Zawel and Reinberg, 1995).

As important as it is, initiation is only the beginning of transcription. After finding the correct place to start (promoter binding) and beginning to polymerize a chain of RNA (initiation), the polymerase must traverse the template (elongation). During this stage, it must pass many obstacles, the first of which is the structure of the template itself. The bacterial chromosome of *E. coli* is packaged into a nucleoid, a structure not yet well understood but utilizing the histone-like HU protein (Drlica, 1987). In eukaryotic cells, the DNA is packaged into nucleosomes and higher order chromatin structures by histones and chromatin associated proteins (Gilmour et al., 1994; van Holde, 1993). If the polymerase loses its hold on either the template or the nascent RNA, the transcript cannot be completed, and the polymerase must begin anew. When it reaches the end of the gene, it must release the RNA and disengage from the template (termination).

All four stages of transcription, 1) promoter binding, 2) initiation and promoter escape, 3) transcript elongation, and 4) termination, are potential targets for regulation (Chamberlin, 1994). The regulation of initiation is quite important, but modulation of either of the two subsequent processes will alter gene expression as well. In the polycistronic operons of prokaryotes, failure to terminate may allow inappropriate expression of a downstream gene. The

attenuation of operons utilizes control of termination as a regulatory strategy. Many amino acid biosynthetic operons are expressed only when the reduced concentration of the appropriate charged tRNA reduces the level of translation. Alternative transcript structures form that allow continued transcription through the operons (Yanofsky, 1988). For the monocistronic transcription units typical in eukaryotes, failure to terminate can occlude a downstream promoter and prevent transcription of the downstream gene (Cullen et al., 1984; Proudfoot, 1986). Failure to complete a transcript will prevent expression of the gene. Errors of inappropriate expression or of failure in expression are potentially lethal for the cell, depending on the gene product.

If much regulation of transcription occurs at initiation, why should the cell regulate the elongation of transcription? In many ways, it makes sense for a cell to regulate this process. To be able to load several polymerase molecules onto a gene such that they are ready to go at a moment's notice allows the cell to produce a gene product it needs very quickly. To be able to terminate the transcription of a transcript in progress if the product becomes unwanted or if another gene product becomes essential allows the cell to save resources and divert polymerase molecules to the locations where they are needed most (Kane, 1994).

Regulation of Transcription Elongation in Prokaryotes

It has long been known that the regulation of elongation is important in prokaryotes. A classical example is the regulation of the tryptophan operon in *E. coli* by attenuation of transcription (Yanofsky, 1988; Yanofsky and Crawford, 1987). A key element of this form of attenuation is that the ribosome begins

translating the mRNA during its synthesis by RNA polymerase. The initial leader region of the transcript encodes a leader peptide containing several tryptophan codons and contains a sequence that causes RNA polymerase to pause. When tryptophan is abundant, the ribosome follows the polymerase closely and a transcription terminating structure composed of a stem-loop and a run of uridine residues forms in the RNA transcript, preventing production of a complete mRNA. A lack of charged tRNA for tryptophan stalls the ribosome translating the leader peptide and allows formation of an antitermination stem-loop in the mRNA, a conformation mutually exclusive with the terminator structure. This allows transcription of the entire operon so that the needed biosynthetic enzymes can be produced. Many amino acid biosynthetic operons in *E. coli* and other enteric bacteria are regulated by attenuation mechanisms similar to that of the trp operon (Yanofsky, 1988).

Many other operons in bacteria are controlled by antitermination mechanisms unlinked to translation. In these cases, it is thought that there are regulatory products other than charged tRNAs that control the formation of antitermination versus termination structures in the RNA (Yanofsky, 1988; Landick and Yanofsky, 1987). Clearly, operons for the ribosomal RNAs must have some method to protect them from rho-dependent antitermination (Adhya and Gottesman, 1978) in the absence of translation. The polymerase is thought to be modified early in the transcription unit to be resistant to termination, similarly to the manner established for antitermination of the λ genome (Holden and Morgan, 1984; Li et al., 1984; Morgan, 1986; Sharrock et al., 1985). The nus proteins, a group of cellular proteins originally identified as essential for N-dependent antitermination within the λ genome, are required for antitermination of the

ribosomal operons (Roberts, 1993; Squires et al., 1993) and possibly protein coding operons such as tnaAB as well (Stewart and Yanofsky, 1985). Most of the nus proteins are essential for viability of *E. coli*, but mutations in the rho protein that diminish rho-dependent termination allow growth of a nusA null mutant (Zheng and Friedman, 1994).

The best characterized system utilizing the regulation of transcriptional elongation occurs during lytic infection of *E. coli* by λ phage. In order to express parts of its genome, λ phage requires the host-encoded antitermination factors, nusA, nusB, nusG, and the ribosomal protein S10 (also called nusE). In the absence of the antitermination proteins, transcription of the λ genome is not completed. Failure to transcribe the entire genome prevents the production of new phage particles (Friedman et al., 1987). Antitermination of the early operons required for lytic function also requires the phage-encoded RNA-binding protein N and an N-utilization (nut) site in the DNA. In order to be functional, the nut site must be transcribed into RNA. The nut RNA sequence, part of which forms a stem-loop structure, allows the formation of a multi-protein complex containing N, nusA, nusG, nusB, S10, and the RNA polymerase. For antitermination throughout the λ genome, it is proposed that the RNA transcript forms a loop such that the nut RNA remains a structural component of the antitermination complex. Although the exact mechanism by which this complex renders the polymerase resistant to termination is not yet clear, this model is one paradigm for the study of the regulation of transcription elongation (Greenblatt et al., 1993; Roberts, 1993).

Expression of the late genes of the λ genome requires nusA, the phage-encoded Q protein, and the Q-utilization (qut) site located in the pR' promoter of the late gene operon (Yang et al., 1987). This promoter is constitutively transcribed

in vivo, but in the absence of the Q protein the RNA is terminated 194 nucleotides downstream at t_R'. This termination precludes production of the proteins required for cell lysis. In the presence of Q, the entire 23 kb late transcript is produced (Friedman et al., 1987). *In vitro*, the polymerase stalls 16 or 17 nucleotides downstream of the start site. This pause, which is specified by interactions between the DNA sequence of the non-template strand of the transcription bubble and the RNA polymerase, is an essential step in Q-dependent antitermination. The stalled polymerase complex is somehow modified and becomes resistant to downstream termination signals (Ring and Roberts, 1994). Q protein binds DNA upstream of the start site and interacts with the RNA polymerase; this interaction changes the footprint of the complex on DNA, shortens the length of the pause, and renders the polymerase resistant to pausing and termination downstream (Yarnell and Roberts, 1992). How the Q protein alters the elongation properties of the polymerase is not understood, but mutations which eliminate the promoter proximal pause also prevent Q function. It is thought that Q may interact specifically with a paused conformation of the polymerase (Ring and Roberts, 1994). This model of Q-mediated antitermination is another paradigm for the regulation of transcriptional elongation; the promoter proximal pauses discovered in many eukaryotic genes (Rougvie and Lis, 1990; Mirkovitch and Darnell, 1992; Krumm et al., 1992, 1993, 1995; Cullen, 1993) bear a striking similarity to the promoter proximal pause in the λ pR' promoter.

Regulation of Transcription Elongation in Eukaryotes

In the last several years, many laboratories have accumulated evidence that transcription elongation is a regulated step within eukaryotic cells. Many viruses of

eukaryotic cells regulate the expression of their genomes through control of transcriptional elongation, including simian virus 40 (SV40) (Hay et al., 1982; Hay and Aloni, 1984), adenovirus (Maderious and Chen-Kiang, 1984; Mok et al., 1984), polyomavirus (Skarnes et al., 1988), minute virus of mice (Resnekov and Aloni, 1989), hepatitis (Guo et al., 1991), bovine papilloma virus (Baker and Noe, 1989) and the human immunodeficiency virus, HIV-1 (Kao et al., 1987; Cullen, 1993). In many cases, study of these viral systems has provided insight into the normal cellular mechanisms to control transcriptional elongation. Many cellular genes in eukaryotes also contain blocks to elongation which appear to be an important component of the regulation of the gene's expression *in vivo* (see Table 1). Blocks to transcriptional elongation have been reported in a variety of genes encoding transcription factors, receptor tyrosine kinases, essential enzymes, and structural components of the cell.

Importance of Regulating Transcriptional Elongation in Eukaryotes

In many cases, the regulation of a given gene was of interest because improper regulation resulted in cancers or diseases in humans (Higgs et al., 1983; Orkin et al., 1985; Bentley and Groudine, 1986, 1988; Bender et al., 1987; Haley and Waterfield, 1991; Chinsky et al., 1989; Hirschhorn, 1992; Shor et al., 1995; McCann and Pegg, 1992). Proper expression of these genes is critical for normal cellular growth and metabolism. Thus, expression of these genes is also controlled by many post-transcriptional regulatory mechanisms, including mRNA stability, translational efficiency, post-translational modifications, protein stability, and enzyme activity. However, recognition of a block to elongation within a transcription unit acts at an earlier step by preventing formation of the mRNA.

Understanding how transcriptional elongation is regulated is essential to fully understand how the expression of these genes is controlled. A brief overview of genes containing blocks to elongation and their cellular functions underscores the importance of this regulatory mechanism.

Many of the first genes characterized as containing blocks to elongation were originally identified as cellular homologues of viral oncogenes, e.g. c-*myc* (Linial et al., 1985; Bentley and Groudine, 1986; Eick and Bornkamm, 1986; Nepveu and Marcu, 1986; Nepveu et al., 1987), c-*myb* (Bender et al., 1987), and c-*fos* (Greenberg and Ziff, 1984; Fort et al., 1987). In many cases, inappropriate expression of these cellular proto-oncogenes is found in association with certain cancers in humans or other vertebrates. Although the normal cellular functions of these genes are still not completely understood, many of them are important in the regulation of cellular proliferation and differentiation.

Some of these cellular proto-oncogenes are nuclear transcription factors activated in response to signals stimulating cellular proliferation or differentiation; others are receptor tyrosine kinases involved in the signal transduction pathways mediating cellular responses to growth factors or hormones. For all of these genes, a regulated block to elongation early in the transcription unit is proposed to be a mechanism which allows rapid modulation of the production of the mRNA (Spencer and Groudine, 1990; Kerppola and Kane, 1991; Krumm et al., 1993; Kane, 1994). Expression of the c-*myc* gene is frequently reactivated in Burkitt's lymphoma. This basic-helix-loop-helix-leucine zipper protein is thought to play a crucial role in the regulation of cell proliferation and differentiation (Marcu et al., 1992; Lüscher and Eisenman, 1990a). L-*myc* and N-*myc* are related to c-*myc* in structure and function but have more restricted patterns of expression in

mammalian tissues. Inappropriate expression of each gene is associated with specific carcinomas (Xu et al., 1991; Krystal et al., 1988). The c-*myb* gene is important in the differentiation of hematopoietic lineages and is often deregulated in malignancies of these cells (Bender et al., 1987; Watson, 1988; Lüscher and Eisenman, 1990b; Toth et al., 1995). Overexpression of the epidermal growth factor (EGF) receptor proto-oncogene (c-*erbB*), a receptor tyrosine kinase, is often associated with human carcinomas (Ozanne et al., 1986; Haley and Waterfield, 1991). The c-*fms* gene, expressed only in mononuclear phagocytes and in placental trophoblasts, also encodes a ligand-dependent protein tyrosine kinase, and mediates the proliferation and differentiation of these cell types in response to macrophage colony-stimulating factor 1, CSF-1 (Yue et al., 1993; Ross et al., 1994). The c-fos protein forms heterodimers with the c-jun protein that are important for transcriptional regulation of many other genes (Glover and Harrison, 1995; O'Shea et al., 1992; Hsu et al., 1992), and c-*fos* transcriptional elongation is regulated in part by calcium ion fluxes (Collart et al., 1991; Werlen et al., 1993; Lee and Gilman, 1994).

Other genes containing blocks to transcriptional elongation encode essential cellular enzymes. These enzymes are intimately involved in the chemical complexities inherent in cells. As with the cellular proto-oncogenes described above, these genes are tightly regulated by a variety of mechanisms in addition to the block to transcriptional elongation. These blocks to elongation are regulated in cell cycle or tissue-specific manners.

For example, adenosine deaminase, a purine metabolic enzyme, catalyzes the irreversible deamination of adenosine or deoxyadenosine to inosine or deoxyinosine. This reaction regulates the cellular concentrations of adenosine and

deoxyadenosine, metabolites with pleiotropic effects (Chinsky et al., 1989). The lack of this enzyme causes severe combined immune deficiency (SCID), a lethal disease when untreated but treatable in some cases by gene therapy (Hirschhorn, 1992). Tissue specific expression of this enzyme is controlled in part by a block to elongation localized to the exon 1/intron 1 region (Chinsky et al., 1989; Chen et al., 1990; Kash and Kellems, 1994), and this block is modulated by TFIIS *in vitro* (Kash and Kellems, 1993).

Another gene proposed to contain a regulated block to elongation encodes the enzyme ornithine decarboxylase (ODC) (Shor et al., 1995), which catalyzes the conversion of ornithine to putrescene, the rate limiting step in polyamine synthesis. ODC activity and polyamines are essential for the proliferation of mammalian cells. Polyamines are reported to influence both the rate and fidelity of protein synthesis (reviewed in Pegg, 1986). Hypusine modification (on a lysine by an addition from the polyamine spermine) of the translation initiation factor eIF-5A, the only protein known to contain this unusual amino acid, is absolutely conserved and is essential for viability (Schnier et al., 1991). Deregulated expression of ODC has been observed in a variety of malignancies in mammals and the ODC gene itself can transform cells (Auvinen et al., 1992). Expression of this gene is tightly controlled at many levels. This regulation may include two blocks to elongation which function by different mechanisms and which are read-through in response to the elongation factor TFIIS *in vitro* (Shor et al., 1995).

The expression of ribonucleotide reductase, a key enzyme required for DNA synthesis, is regulated by a block to elongation. This enzyme catalyzes the direct reduction of ribonucleotides to the corresponding deoxyribonucleotides. The mammalian enzyme is composed of two subunits, R1 and R2. The level of R1

protein is constant throughout the cell cycle. R2 protein levels are limiting and determine enzymatic activity. Both mRNAs are specifically expressed in S-phase. The increase in R2 mRNA at this point in the cell cycle is due in part to a release from a block to elongation located in the exon 1/intron 1 boundary, near a potential ME1a1 binding site for the MAZ protein (Björklund et al., 1992), which will be discussed in further below.

Regulation of transcriptional elongation appears to play a role in production of the erythroid specific isozyme of porphobilinogen deaminase. This gene encodes the third enzyme in the biosynthetic pathway for heme, an essential prosthetic group in numerous proteins including hemoglobin and mitochondrial cytochromes. The gene is expressed in all tissues from either of two promoters and produces two isozymes which differ at the amino-terminus. The first promoter, P1, is the housekeeping promoter; the second promoter, P2, is erythroid specific. When P2 becomes activated in erythropoietic cells, use of P1 also increases but most of these transcripts now stop in the first intron, upstream of P2 (Beaumont et al., 1989). This block to elongation for P1 initiated transcripts may be involved in preventing promoter occlusion of P2 (Porcher et al., 1995).

Some genes which encode "structural" components of cells are also regulated at the level of transcriptional elongation. The human histone H3.3 gene contains a block to elongation in the first intron (Reines et al., 1987) which is regulated in a cell cycle dependent manner (D. Wells, personal communication). This block is read through in response to TFIIS *in vitro* (Reines et al., 1989; SivaRaman et al., 1990). The α-tubulin gene of Xenopus also contains a block to elongation (Middleton and Morgan, 1989; Hair and Morgan, 1993).

A variety of other genes are also regulated during transcriptional elongation. The chicken β^A globin gene may contain a block within the gene (Lois et al., 1990). Expression of two retinoic acid receptor (RAR) isoforms is induced by a retinoic acid-dependent release of a block to elongation (Mendelsohn et al., 1994). Induction of expression of tumor necrosis factor α (TNF-α) by lipopolysaccharides has recently been reported to be due to the release of a block to elongation (Biragyn and Nedospasov, 1995). Developmental regulation of the bovine opsin gene may involve control of transcriptional elongation (DesJardin et al., 1995). Recently, it has been reported that the mammalian genes encoding the Factor VIII clotting factor and the cystic fibrosis transmembrane conductance regulator (CFTR) both contain blocks to elongation. For both of these genes, the block in transcriptional elongation may be a factor in the low levels of expression from vectors intended for gene therapy (Koeberl et al., 1995).

In many cases, these blocks to transcriptional elongation appear to be essential facets of the regulation of the gene's expression. Proper control of these blocks to elongation is required to insure that the gene is expressed only in the correct tissues with the appropriate timing in the cell cycle and the development of an organism. The regulation of transcriptional elongation is a mechanism that allows the cell to fine tune the expression of a given gene. Proper regulation of the genes described in the preceding pages is critical for normal cellular growth and development. However, the mechanisms by which transcription elongation is regulated in eukaryotic cells are poorly understood.

Locations of Blocks to Elongation

The reported blocks to elongation in genes transcribed by RNA polymerase II can be divided into three basic types based on location within the gene: 1) promoter proximal pauses, 2) blocks to elongation near the first exon/intron junction, and 3) blocks to elongation downstream of the polyadenylation [poly(A)] site. Termination at the blocks represented in the first two categories would result in premature termination of the transcript. Termination at the blocks represented in the third category is considered the appropriate event, but failure to do so produces continued elongation and the potential for occlusion of downstream promoters. Therefore, these blocks will be included in this discussion on the regulation of transcriptional elongation. At least two genes, c-*myc* and c-*fos*, have been reported to contain both a promoter proximal block and a block within the first intron (Krumm et al, 1992, 1993, 1995; Strobl and Eick, 1992; Plet et al., 1995). Further examination of the two genes, especially c-*myc*, suggests that the distinction between promoter proximal pauses and blocks located further downstream within the transcription unit may be artifactual; this will be discussed further below. Blocks to elongation, whether early in the transcription unit or downstream of the poly(A) site, can have important regulatory consequences. Table 1 presents a list of cellular genes which have been reported to contain a block to elongation and the reported positions of these blocks within the transcription unit.

Techniques for the Identification of Blocks to Elongation

While the list of identified genes potentially regulated during the transcription elongation step is not long, it may represent only the tip of an

iceberg. The identification of an *in vivo* block to elongation during the process of transcription is not a trivial task. A Northern blot provides information only on steady state levels of an mRNA; only stable species are analyzed. The techniques which allow detection of partially completed RNA species are more difficult due to the instability of unprocessed, non-polyadenylated transcripts. There have been three major techniques used for the identification of blocks to transcriptional elongation: 1) nuclear run-on, 2) nuclease protection of RNAs transcribed in Xenopus oocytes, and 3) *in vivo* footprinting. A basic overview of these techniques provides the basis for what is known about blocks to elongation within transcription units and also provides an explanation for why many details remain controversial or poorly understood.

Nuclear run-on (alternatively called nuclear run-off) analysis (Weber et al., 1977) has been used to detect many of the reported blocks to elongation. This technique requires isolation of nuclei such that engaged RNA polymerase II molecules remain on the template in a complex containing the nascent transcript. Radiolabelled nucleotides are added back under conditions that prevent *de novo* initiation of transcription. This treatment allows polymerase molecules to add an additional 100-200 nucleosides to previously initiated RNA transcripts. These RNAs are isolated and hybridized to several, immobilized DNA probes from the gene of interest. Done properly, this technique provides a view of the relative density of RNA polymerase II molecules along a transcription unit at the time the nuclei were isolated. A gene containing a block to elongation is characterized by a significant decrease in hybridization signal to the 3'-region of the gene relative to the hybridization signal to the 5'-region of the gene.

However, several technical aspects of the use of nuclear run-on to analyze *in vivo* transcription have made it difficult to interpret the results obtained with this technique. Resolution is limited by the size of the probes required for efficient hybridization and also by the requirement to label the transcripts by increasing their lengths. Improper selection of probes for nuclear run-on can produce erroneous results as well. When double stranded DNA probes are used, some of the signal observed might be produced by transcription of the non-template strand (with respect to the gene of interest). The presence of repeated sequences in a probe also can produce artificially high hybridization signals. The amount of time the nuclei are incubated in the presence of radiolabelled nucleotides can also affect the relative intensities of hybridization to the probes used. In addition, attempts to determine whether a gene is transcribed in a given cell line or condition have sometimes used the complete cDNA of the gene as a nuclear run-on probe. Positive signals have been interpreted to mean the gene is transcribed (Blanchard et al., 1985; Dony et al., 1985). However, nuclear run-on done in this manner does not distinguish between complete or partial transcription of a gene. Despite the difficulties, the use of a series of relatively small, single stranded probes that span the gene of interest can produce a reasonably good representation of the polymerase densities in various regions of a gene (Eick et al., 1994).

Attempts to determine the 3'-ends of transcripts produced *in vivo* have been unsuccessful. The low resolution of nuclear run-on as typically performed with fairly large probes does not allow precise determination of the location of the block to elongation. Prematurely terminated transcripts are presumed to be highly unstable in cells due to the lack of a poly(A) tail or other protective structure. As

will be discussed further below, the location of the poly(A) tail in a eukaryotic mRNA results from processing and does not represent the location where RNA polymerase II terminates transcription.

Xenopus oocytes have been utilized to determine the location of the primary 3'-ends of RNAs directly since many non-polyadenylated RNAs are stable in these cells (Green et al., 1983). When templates containing a mammalian gene of interest are injected into oocytes, transcription in the frog oocyte often starts at the same positions observed during transcription in mammalian cells. Furthermore, in come cases transcription in Xenopus appears to faithfully duplicate the locations of blocks to elongation in mammalian genes (Nishikura, 1986; Bentley and Groudine, 1988). It is then possible to isolate RNAs from injected oocytes and use S1 nuclease to determine the actual 3'-ends of RNAs produced when the polymerase recognized a block to elongation. This technique has allowed determination of the specific sites of some blocks to elongation. Results obtained in this manner often corroborate the estimates obtained with nuclear run-on in mammalian cells.

However, there are problems associated with characterization of blocks to elongation by transcription in Xenopus oocyte as well. Injection of too much template DNA can increase the amount of premature termination observed and co-injection of control DNAs can alter the results obtained on the experimental template (Spencer and Kilvert, 1993; Middleton and Morgan, 1990; Meulia et al., 1993). This effect has been interpreted to mean that Xenopus oocytes contain an elongation factor that becomes limiting as more template is used. A more serious problem is that transcription of a mammalian c-*myc* template in these oocytes does not produce the promoter proximally paused polymerases observed in mammalian

cells on the same template (Meulia et al., 1993). The 3'-ends seen when a c-*myc* template is transcribed in Xenopus localize to the exon 1/intron 1 boundary (Bentley and Groudine, 1988). This difference between transcription in Xenopus and mammalian cells limits the usefulness of the Xenopus system for the study of blocks to elongation which occur in mammalian genes.

For the majority of genes reported to contain a block to transcriptional elongation in the first intron, the block was characterized by either or both of the two techniques described above. In many cases, the nuclear run-on detection was performed with fairly large probes, and the specific location of the block to elongation was determined through S1 nuclease mapping or RNase protection of RNAs produced by transcription in Xenopus oocytes. For the technical reasons discussed above, the majority of genes reported to contain blocks to elongation in this position will need to be reevaluated to determine the specific location of the block to elongation; the presence of a block to elongation is not in question. There are many genes which produce equivalent hybridization signals when the 5' and 3'-regions of the gene are used as probes for nuclear run-on (Rougvie and Lis, 1990). Thus a decrease of polymerase density in the 3'-region of a gene as determined by nuclear run-on does indicate a block to elongation, although the nature and location of the block may not be determined.

A third, more recently developed technique for characterizing stalled polymerases, is *in vivo* $KMnO_4$ footprinting, which utilizes the preferential reactivity of $KMnO_4$ towards unstacked thymidine residues relative to its reactivity with stacked, paired thymidine residues in double stranded DNA. Thus, this reagent allows detection of RNA polymerase open complexes because the thymidine residues within the single stranded "transcription bubble" region are

preferentially attacked by this reagent which penetrates intact cells. This technique demonstrated DNA melting *in vivo* by *E. coli* RNA polymerase (Sasse-Dwight and Gralla, 1988, 1989, 1990). Effective use of this reagent for *in vivo* footprinting of eukaryotic polymerase complexes requires sensitive methods of sequencing genomic DNA, which have only been developed recently (Mirkovitch and Darnell, 1991, 1992; Mirkovitch et al., 1992). With this technique, it is possible to visualize the position of RNA polymerase molecules on a gene *in vivo* when the rate-limiting step in transcription occurs after open complex formation, such that the open complexes are relatively long lived *in vivo*.

In vivo footprinting is ideal for detecting genes with a block to elongation in a specific location. On transcribed genes without transcriptional blocks, the polymerases are located heterogeneously along the template. When there is a block, a large percentage of cells within a population contain a polymerase paused at that position, generating a single stranded region detectable with $KMnO_4$ footprinting. A significant advantage of this technique over nuclear run-on or nuclease mapping of RNAs from transcription in Xenopus oocytes is the determination of the actual location of the paused polymerase. Increased interest in detection of paused complexes is spurring improvements in the sensitivity of *in vivo* footprinting techniques (Rasmussen and Lis, 1995).

These three techniques have been the major methods used to identify and characterize transcriptional blocks to elongation. It is clear from the examples described above that the regulation of transcriptional elongation is a ubiquitous regulatory mechanism. The mechanisms which determine these blocks to transcriptional elongation are not yet understood. The following sections will

explore what is currently known about the determinants of transcriptional blocks to elongation.

Blocks to Elongation in the 3'-Regions of Genes

The most distal type of block to elongation, occurring downstream of the poly(A) site, was the first type to be identified in a eukaryote. Genes determined to contain blocks to elongation in the 3' region of the gene include β-globin (Hofer and Darnell, 1981), α-globin (Weintraub et al., 1981), α-amylase (Hagenbüchle et al., 1984), ovalbumin (LeMeur et al., 1984), dihydrofolate reductase (Frayne et al., 1984), calcitonin (Amara et al., 1984), gastrin (Sato et al., 1986), histone H2A (Johnson et al., 1986), both U1 and U2 snRNAs (Hernandez and Wiener, 1986; Neuman de Vegvar et al., 1986), human complement C2 (Ashfield et al., 1991, 1994), and immunoglobulin μ (Galli et al., 1987). The blocks to elongation in the majority of these genes appear to be involved in appropriate termination downstream of the poly(A) site. Contrary to initial hypotheses, RNA polymerase II does not just fall off the template DNA at random after it has transcribed the poly(A) site; the termination of transcription at the end of an RNA polymerase II transcription unit is an active process requiring the processing machinery that produces the mature 3'-end of the transcript (Proudfoot, 1989).

The requirement for 3'-end processing in transcriptional termination has been clearly demonstrated for several genes. Transcription of the globin genes normally terminates heterogeneously downstream of the poly(A) site, as shown by nuclear run-on assays (Citron et al., 1984), in a 1 kb region that contains A-T rich sequence elements which may be signals for the polymerase to pause. However, the region downstream of the poly(A) site alone is insufficient to cause termination

of transcription; the poly(A) site and processing machinery are also required for termination (Falck-Pederson et al., 1985; Reines et al., 1987). Mutations in the poly(A) signal which prevent polyadenylation cause increased transcription past the normal termination region (Whitelaw and Proudfoot, 1986; Connelly and Manley, 1988). It is thought that two elements, a functional poly(A) site and a theoretical downstream pause site, are required for coupling of transcriptional termination to the processing which produces the mature, polyadenylated 3'-end. The mechanism of this coupling is not yet elucidated. Transcriptional termination of non-polyadenylated RNA polymerase II transcripts, such as histone mRNAs, and the U1 and U2 snRNAs, also requires active 3'-end processing mechanisms (Birnstiel et al., 1985; Hernandez and Weiner, 1986; Neuman de Vegvar et al., 1986; Hernandez and Lucito, 1988).

It might seem that failure to terminate properly after the entire coding sequence of the gene has been transcribed should not be detrimental to the expression of that gene. In a prokaryote where the message is translated as it is transcribed, this might be true. However, in a eukaryotic cell the message must be completed and processed before it is sent to the cytoplasm to be translated. It is clear that failure to produce the proper 3'-end of the globin genes can be deleterious. Mutations that destroy the normal poly(A) sites reduce globin mRNA and protein levels and cause mild thalassaemias (Higgs et al., 1983; Orkin et al., 1985). Similarly, a deletion which removes the 3'-end formation signal of the *S. cerevisiae CYC1* gene causes the formation of aberrantly long transcripts, some increased in length by more than 2000 nucleotides, and reduces its expression by 90% (Guo et al., 1995; Zaret and Sherman, 1982). The blocks to elongation in the 3' regions of the majority of genes listed in section C of Table 1 may function

similarly to those in the globin genes to cause transcriptional termination downstream of the poly(A) site.

An additional possible consequence of failing to terminate transcription at the end of a gene is occlusion of the promoter of a downstream gene (Cullen et al., 1984; Proudfoot, 1986). Transcriptional interference has been shown to occur in a manipulated system composed of duplicated α-globin genes (Proudfoot, 1986). This phenomenon may be relevant *in vivo* as well. A consensus transcription termination sequence in the promoter of the *S. cerevisiae TRP1* gene is required for efficient transcription of the gene (Braus et al., 1988).

Proper termination between the human complement C2 and Factor B genes, thought to be important for normal expression of both genes, may be mediated by a termination factor. The coding region of the human Factor B gene begins only 421 base pairs downstream of the poly(A) site for the complement C2 gene and is expressed at a higher level than the upstream gene. Termination of transcription occurs very shortly after the poly(A) site of the complement C2 gene and requires the 160 base pair region downstream of the poly(A) site. The intergenic region, conserved in mammalian species (Moreira et al., 1995), contains a binding site for the MAZ (c-myc-associated zinc finger) protein. Binding by MAZ bends the DNA and mutations of the MAZ site that prevent binding eliminate termination. However the MAZ binding site is not sufficient to cause termination; the surrounding DNA sequences are also required (Ashfield et al., 1991, 1994). In addition to the MAZ sites in the C2-Factor B intergenic region and the previously defined sites in the c-*myc* promoter region (Miller et al., 1989; Pyrc et al., 1992; Bossone et al., 1992), MAZ sites have been identified between the closely spaced human genes g11 and C4, within an intron of the murine IgM-D gene (Ashfield et

al., 1994), and in the first intron of the gene for the ribonucleotide reductase R2 subunit (Björklund et al., 1992). It is postulated that the MAZ protein may be a termination factor that makes specific contacts with RNA polymerase II (Ashfield et al., 1994).

Regulated blocks to elongation may play key roles in developmental regulation of immunoglobulin proteins. In the immunoglobulin μ heavy chain gene, a block to elongation in the 3' region may be involved in the regulation of the production of membrane bound, μ_m, and secreted, μ_s, forms of the protein. The mature B cell produces both forms of the protein. When antigen stimulates proliferation and terminal differentiation of the B cell, only the secreted form is produced. Both forms are encoded by one gene; two variant mRNAs are produced that differ in the presence of distinct 3' exons (Galli et al., 1987). Several factors contribute to control of which mRNA species is produced. The μ_s form is encoded by the shorter mRNA, produced when the weaker upstream poly(A) site is used. In mature B cells, this weak poly(A) site is used inefficiently. When transcription continues past the second, downstream poly(A) site, the μ_m mRNA is produced instead. The switch to production of only the secreted form of μ correlates with an increase in transcriptional termination between the weak and the strong poly(A) site (Galli et al., 1987). It appears that B cells contain an activity that destabilizes the formation of the processing complex on the μ_s poly(A) site (Yan et al., 1995). Inefficient use of this first poly(A) site allows RNA polymerase II to continue transcription through to the μ_m poly(A) site. The switch from expression of IgD to IgM from the IgM-D locus is also regulated by transcriptional termination in an intron 5' of the δ exons (Mather et al., 1984; Yuan and Tucker, 1984). The role of a putative binding site for the MAZ protein, recently been found in this region of the

mouse gene (Ashfield et al., 1994), in the regulation of transcriptional termination in this region remains to be demonstrated.

Blocks to Elongation Within Genes

Eukaryotic cells also regulate the process of elongation by RNA polymerase II within transcription units. The regulation of transcriptional elongation was observed in transcription units of viruses which infect eukaryotic cells before it was identified in transcription units of eukaryotic cells (Hay et al, 1982; Maderious and Chen-Kiang; 1984). The study of this process in viral gene expression has significantly advanced our understanding of this type of regulation (Jones and Peterlin, 1994). The c-*myc* gene was the first example discovered of a cellular, eukaryotic gene containing a regulated block to transcriptional elongation within the transcription unit (Bentley and Groudine, 1986; Nepveu et al., 1986; Eick and Bornkamm, 1986). Although the c-*myc* gene contains one of the most studied blocks to elongation in a eukaryotic gene, the nature of the block to elongation and the mechanism of its regulation are still not clear. A brief chronology of the study of the mechanisms that regulate the block to elongation in the mammalian c-*myc* gene illustrates the complexity of studying these reactions in cells.

Nuclear run-on analysis of mammalian c-*myc* genes indicated that the untranslated first exon sequences contained a higher density of RNA polymerase II complexes than the downstream regions of the gene and localized the block to elongation to the exon 1/intron 1 junction. When the steady state levels of c-*myc* mRNA are reduced in cell lines stimulated to undergo terminal differentiation, usage of the block to elongation increases (Bentley and Groudine, 1986; Eick and Bornkamm, 1986; Nepveu and Marcu; 1986). Deregulation of the block to

elongation correlates with reactivation of the gene in Burkitt's lymphoma (Cesarman et al., 1987; Spencer et al., 1990). The block to elongation was localized to two stretches of T residues near the exon 1/intron 1 junction by S1 nuclease mapping of RNAs produced in Xenopus oocytes (Bentley and Groudine, 1988). Sequences that could putatively form RNA structures very similar to the read-through and terminator RNA conformations of bacterial attenuators are present in this region (Eick and Bornkamm, 1986; Miller et al., 1989). The significance of the putative stem-loop structures remains in question since purified RNA polymerase recognizes the runs of T's as a block to elongation *in vitro* but RNA secondary structure is not required (Kerppola and Kane, 1988, 1990; Dedrick et al., 1987). Further, elimination of the hairpin did not eliminate the block to elongation in HeLa cells (Wright and Bishop, 1989). Sequences from the first exon of the c-*myc* gene were shown to be sufficient to cause premature termination of transcription when placed in the correct orientation downstream of several other promoters (Bentley and Groudine, 1988; Wright and Bishop, 1989; Roberts and Bentley, 1992). Alone these results could suggest a simple explanation: sequences in the first exon are necessary and sufficient to induce RNA polymerase II to arrest transcription and this block to elongation is regulated when cells differentiate.

Additional levels of complexity are present though. Usage of the block to elongation in exon 1 has been observed to depend on which of the two c-*myc* promoters was used for initiation (Spencer et al., 1990). For both the human and murine c-*myc* genes, the major promoter P2 is 160 base pairs downstream of the P1 promoter and produces about 80% of steady state c-*myc* mRNAs in normal cells. In c-*myc* genes in Burkitt's lymphomas, a switch to primary usage of the P1 promoter is frequently observed. Analysis of usage of the block to elongation

during transcription in Xenopus oocytes on mammalian templates lacking either the P1 or the P2 promoter indicates that initiation from P1 leads to read-through of the block to elongation while initiation from P2 leads to arrest in exon 1 (Spencer et al., 1990). Deletion mutants analyzed by nuclear run-on in mammalian cells suggested that a sequence element containing the ME1a1 binding site for the MAZ protein in between the P1 and P2 promoters is required for initiation from P2 (Asselin et al., 1989) and for P2-initiated complexes to recognize the block in exon 1, perhaps by modulating the composition of the elongating complex (Miller et al., 1989). Analysis of the P2 promoter indicates that three protein binding sites ME1a2, E2F, and ME1a1 are all required for efficient initiation from P2. The E2F site is required for transactivation but requires either the ME1a2 or ME1a1 sites to function (Moberg et al., 1992).

However, other work determined that P1-initiated transcripts are prematurely terminated 5' of the P2 promoter (Wright et al., 1991; Roberts et al., 1992) and that sequences in the P2 promoter are required for termination of P1-initiated transcripts with a possible role for the ME1a1 binding site in preventing occlusion of the P2 promoter by P1-initiated polymerase complexes (Meulia et al., 1992). This latter hypothesis is consistent with the determination that the *myc*-associated zinc finger (MAZ) protein, which binds both the ME1a2 and ME1a1 sites in the c-*myc* promoter (Pyrc et al., 1992; Bossone et al., 1992), is required for termination downstream of the human complement C2 gene (Ashfield et al., 1991, 1994). A possible reason that P1-initiated complexes had been observed to read-through the block to elongation in exon 1 is that these complexes, which had transcribed past the arrest sites flanking the P2 promoter, were a subset of complexes more resistant to pausing or termination while those P1-initiated

complexes prone to premature termination arrest near the P2 promoter (Roberts and Bentley, 1992). This work confirmed that promoter usage affected the ability of elongation complexes to recognize the downstream block to elongation in exon 1 but the nature of these promoter effects was unclear.

Further work to elucidate how promoter usage controls the elongation properties of RNA polymerase II indicates that the actual block to elongation for polymerases initiated at the P2 promoter is immediately proximal to the promoter. The observance of a block to elongation at the exon 1/intron 1 junction may be artifactual or may represent a second level of elongational control. Kinetic analysis of nuclear run-on assays performed in mammalian cells with very short probes indicates that the block to elongation in the c-*myc* gene is very close to the promoter (Strobl and Eick, 1992; Eick et al., 1994). *In vivo* footprinting with $KMnO_4$ indicates the presence of a polymerase complex paused at +30 with respect to the P2 promoter (Krumm et al., 1992). The use of the detergent Sarkosyl or high salt concentrations in nuclear run-on assays activates polymerases paused proximal to the promoter to continue elongation; these artificially activated polymerases may be responsible for the observed transcription past the block at +30 through the remainder of the exon 1 sequences (Krumm et al., 1992; Eick et al., 1994).

In vivo footprinting of the c-*myc* gene in a Burkitt's lymphoma cell line, where the c-*myc* gene is translocated into the vicinity of an immunoglobulin enhancer and transcribed primarily from the P1 promoter, indicates that the pause proximal to the P2 promoter is no longer observed. Although a model involving promoter occlusion of P2 by increased use of P1 is a possibility, the data do not suggest a change in the level of the P2 promoter usage in Burkitt's lymphoma cells

(Strobl et al., 1993). Thus it may be that the polymerase paused at the P2 promoter normally blocks the elongation of P1-initiated polymerases. The loss of the paused complex at P2, by some interaction with the immunoglobulin enhancer, allows an increase in productive transcription from P1 since these complexes are no longer terminated at the P2 promoter. An alternative possibility is that the observed switch to preferential usage of the P1 promoter in Burkitt's lymphomas is due to the ability of the immunoglobulin enhancers to stimulate the P1 promoter but not the P2 promoter (M. Groudine, personal communication).

The determination that the block to elongation in the c-*myc* gene is determined in the promoter proximal region and that the localization of the block to the exon 1/intron 1 region may be artifactual has serious implications for the interpretation of experiments suggesting the presence of blocks to elongation in the first intron regions of other genes. Nuclear run-on analysis of c-*myc* transcription in Xenopus oocytes with the same probes which allow determination of the promoter proximal block in mammalian cells indicate that the promoter proximal block to elongation is not duplicated in Xenopus oocyte transcription (Meulia et al., 1993). Combined with the observation that the amount of premature termination of transcription in Xenopus is influenced by the amount of template injected (Spencer and Kilvert, 1993), this result indicates that the use of transcription in Xenopus oocytes to determine the sites of premature transcription termination may not, in many cases, reflect the actual location of blocks to elongation in mammalian cells. Despite the suspicion cast on the locations of many blocks to elongation, these results with c-*myc* do not cast doubt on the presence of blocks to elongation in other genes. It has been clearly demonstrated in Drosophila that some genes (hsp70, hsp26, β-tubulin, Gapdh-1, Gapdh-2, and

polyubiquitin) contain promoter proximally paused polymerases while other genes (actin and histone H1) do not (Rougvie and Lis, 1990). As mentioned in the above discussion on nuclear run-on, the unexpectedly high polymerase densities observed in the early regions of some genes do seem to correlate with regulation of a step of the transcription cycle after initiation. Many of these blocks may turn out to be more promoter proximal than first thought. Others may prove to be caused by entirely different mechanisms. For example, the androgen receptor gene is reported to contain a block to elongation that is mediated by two single stranded nucleic acid binding proteins (Grossman and Tindall, 1995).

The discovery that the primary block to elongation for polymerases initiated from the c-*myc* promoter was in the promoter proximal region indicated that the elongation block in this gene was similar to other well characterized blocks to elongation in eukaryotic genes such as the Drosophila hsp70 gene or the long terminal repeat (LTR) of HIV-1. For all three of these promoters, sequences in the promoter and initial transcribed region determine the formation of polymerase complexes that pause shortly after initiation (Krumm et al., 1992, 1993, 1995; Jones and Peterlin, 1994; Lee et al., 1992; Rasmussen and Lis, 1995).

A variety of RNA secondary structures, some of which are analogous to those of bacterial terminators, have been proposed by many investigators to play a role in causing eukaryotic RNA polymerases to pause or arrest transcription. The promoter proximal pause at +30 in the c-*myc* gene is just downstream of a potential G-C rich stem loop structure and a run of T residues although this putative structure has not been demonstrated to play a role in causing the polymerase to pause at this position (Krumm et al., 1992, 1993, 1995). Previously, similar RNA secondary structures have been proposed in the sequences of HIV-1,

adenovirus, SV40, and in the exon 1/intron 1 sequences of c-*myc* and several other genes; roles for these structures in causing blocks to elongation have also been proposed (Eick and Bornkamm, 1986; Bengal and Aloni, 1991; Kessler et al., 1989; Resnekov et al., 1989).

Although the putative formation of stem-loop structures by the nascent RNA transcript in the vicinity of blocks to elongation in eukaryotic genes is frequently reported, this type of structural element has never been demonstrated to play any role in causing RNA polymerase II to pause or arrest transcription. Indeed there is evidence indicating the these structures are not important, at least for some genes (Kerppola and Kane, 1990; Wiest and Hawley, 1990; Hair and Morgan, 1993). For transcription from the HIV-1 LTR promoter, it has been demonstrated that the TAR RNA stem loop structure is required only for transactivation by the tat protein and is not required for the production of short transcripts from the HIV-1 LTR promoter. The fusion of the inducer of short transcripts (IST) element downstream of other promoters is sufficient to confer the production of short transcripts (Ratnasabapathy et al., 1990; Sheldon et al., 1993; reviewed in Jones and Peterlin, 1994).

The Effects of Initiation Upon Elongation

Recently it has become apparent that the initiation reaction plays a significant role in determining the elongation properties of RNA polymerase. Some promoters tend to set up transcription initiation complexes that are incompetent to pass through pauses or blocks to elongation downstream, while others dictate the formation of elongation competent polymerase complexes. For the mammalian c-*myc* gene, the Drosophila hsp70 gene, and the HIV-1 LTR, sequences immediately

upstream and downstream of the start site seem to play a role in the formation of complexes that pause just downstream of the promoter (Krumm et al., 1992, 1993, 1995; Lee et al., 1992; Lu et al., 1993). The sites of blocks to elongation observed by transcription in Xenopus oocytes may reflect the inability of these elongation "incompetent" polymerases to traverse sequences which may cause polymerases to pause (M. Groudine, personal communication). For the HIV-1 LTR promoter, the IST element is required for the formation of complexes that produce transcripts only 55-60 nucleotides in length (Jones and Peterlin, 1994).

How these promoters dictate the formation of elongation incompetent complexes is not understood. Possibly DNA sequences interact with the RNA polymerase directly to cause these promoter proximal pauses similarly to the sequence determinants for the pause at +16 and +17 of the λ pR' promoter (Ring and Roberts, 1994). Comparison of the initial sequences of the Drosophila genes hsp70, hsp26, and β-tubulin, which all contain paused polymerases *in vivo* as determined by $KMnO_4$ footprinting, indicates some similarities. Although the functional significance remains to be determined, the region of sequence similarity is positioned within the transcription bubble of the paused polymerase (Giardina et al., 1992). Alternatively, protein factors may be the determinants of promoter proximal pausing. In the case of the hsp70 gene, the binding of the GAGA factor at sites upstream of the start site is required for the formation of paused transcription complexes while the heat shock factor (HSF) is involved in release of the paused complexes upon transcriptional induction by heat shock (Lee et al., 1992). The promoter proximal pausing may be required for modification of the polymerase complex to an elongation competent form.

To understand the formation of promoter proximal paused complexes will require elucidation of the transcription initiation reaction. A great deal has been learned in the last few years about the complex protein:protein interactions in the RNA polymerase II pre-initiation complex. *In vitro* studies have revealed an ordered assembly of the pre-initiation complex. The first factors, the TATA-binding protein (TBP) TFIID complex, bind the promoter sequences and then recruit other factors and eventually the polymerase to the promoter (Conaway and Conaway, 1993; Buratowski, 1994; Tjian and Maniatis, 1994). As the various factors have become more highly purified and the genes for them cloned, these *in vitro* studies have allowed more and more precise determination of the functions of each of the polypeptides present in the basal initiation factors.

These studies on initiation complex formation of eukaryotic RNA polymerase II provide little insight into how the c-*myc* or hsp70 promoters direct the formation of elongation complexes. As yet only a small number of well characterized, strong promoters have been used in studies on initiation complex formation (Aso et al., 1994). There has been little investigation of how various promoters may direct the formation of initiation complexes with different properties, although it is almost certain that this is the case. One of the best characterized examples is the transcription of the U1 and U2 snRNAs by RNA polymerase II. For the correct termination event, these transcription units must be initiated from either the U1 or the U2 promoter (Hernandez and Weiner, 1986; Neuman de Vegvar et al., 1986).

Even though most studies on the formation of RNA polymerase II initiation complexes have not directly addressed how a promoter may control the elongation properties of the RNA polymerase, there are a few clues. Studies using

purified general initiation factors have indicated that transcriptional activators act on at least two distinct steps of the initiation reaction. Some transcriptional transactivators function to stimulate the formation of elongation competent complexes while others stimulate the rate of initiation (Yankulov et al., 1994). Some transactivators stimulate both initiation and elongation. The Epstein-Barr virus nuclear antigen 2 (EBNA 2) interacts with TFIIB and with the TATA-binding protein associated factor 40 (TAF40) to promote recycling of TFIIB back to the promoter after it is released by escape of the polymerase from the promoter (Tong et al., 1995a). The EBNA 2 transcriptional activator also interacts with TFIIH to stimulate escape of polymerase from the promoter (Tong et al., 1995b). TAFs in the TBP-containing TFIID complex interact directly with various transcriptional activators and are required for activated but not for basal transcription (Tjian and Maniatis, 1994). The characterization of which of these two steps are stimulated by various transcriptional activators will provide insight into the rate-limiting steps in gene expression.

Promoters which set up elongation incompetent complexes may recruit a different subset of the general initiation factors, transcriptional activators, or elongation factors than those which set up elongation competent complexes. The transcription initiation factors TFIIE and TFIIH have been shown to be required for promoter clearance but not for abortive initiation or for elongation using purified components *in vitro* (Goodrich and Tjian, 1994; Maxon et al., 1994). TFIIH has been shown to contain a kinase activity capable of phosphorylating the carboxyl-terminal domain (CTD) of the largest subunit of RNA polymerase II (Shiekhattar et al., 1995; Drapkin and Reinberg, 1994; Maldonado and Reinberg, 1995). It is quite interesting that this activity is inhibited by 5,6-dichloro-1-β-D-ribofuranosyl-

benzamidazole (DRB) (Yankulov et al., 1995), a compound which has long been known to inhibit "processive" or elongation competent transcription but not "non-processive" or elongation incompetent transcription (Chodosh et al., 1989; Roberts and Bentley, 1992). In contrast, the activities of two elongation factors TFIIS and TFIIF are insensitive to DRB (Kephart et al., 1992).

Although promoter escape for *E. coli* RNA polymerase is currently thought of as the release of the σ subunit and the first translocation, this precise definition may not be appropriate for eukaryotic RNA polymerases which have many more subunits and associated factors. The promoter proximal pauses may be a component of a mechanism to regulate promoter escape. In this type of scenario, the polymerase paused 20 to 40 base pairs downstream of the start site would be stalled for lack of promoter escape factors. The basal promoter might only be sufficient to recruit factors required for the first few steps of transcription, binding of the polymerase to the promoter, the transition from a closed promoter complex to an open promoter complex, and initiation of the RNA transcript. After completing these steps the polymerase would pause. Activation of the promoter would bring the promoter clearance factors to release the polymerase from the promoter proximal region. There is some appeal to this model of regulated promoter clearance. However, there is not yet any evidence that this occurs *in vivo*. In addition, the current data do not address the role of elongation factors in modulating the elongation competence of RNA polymerase II complexes *in vivo*. The recent improvements in techniques for *in vivo* footprinting may provide methods to determine precisely which factors associate with the polymerase at a given position.

Transcription Elongation Factors

At least six different transcription elongation factors for RNA polymerase II have been reported, including TFIIS (S-II, P37), TFIIF, Elongin (SIII), P-TEFa, P-TEFb, and factor 2; each is described below. Another elongation factor activity, TFIIX, has not been further characterized since the initial report (Bengal et al., 1991). The existence of a negative transcription elongation factor, N-TEF, has been hypothesized (Marshall and Price, 1992).

Several different transcription elongation factors have been identified in Drosophila, including TFIIF (originally called factor 5 in the Drosophila system), P-TEFa (Marshall and Price, 1992, 1995), P-TEFb (Marshall and Price, 1992, 1995), and factor 2 (Marshall and Price, 1995). Three of these factors, P-TEFa, P-TEFb, and factor 2, effectively reconstitute efficient DRB-sensitive transcription *in vitro* with the Drosophila system. However, only P-TEFb is absolutely required. Chromatographic properties and subunit composition distinguish P-TEFb from the known basal initiation factors. Although P-TEFb bears no obvious similarity to known kinases involved in transcription, the sensitivity to DRB suggests that it may be a kinase (Marshall and Price, 1995) and like TFIIH, P-TEFb plays a role in conversion of elongation incompetent complexes to elongation competent complexes. Since P-TEFb is probably not Drosophila TFIIH, it is not yet clear whether these two factors regulate the same or different steps in transcription.

The transcription elongation factor TFIIF (human RAP30/74, rat βγ, Drosophila factor 5) is also one of the general initiation factors (Price et al., 1987; Finkelstein et al., 1992; Chang et al., 1993; reviewed in Conaway and Conaway, 1993; Buratowski, 1994). This two-subunit (three in *S. cerevisiae*) factor interacts with TFIIB to promote selective binding of RNA polymerase II to the initiation

complex. This factor binds directly to the polymerase and prevents non-specific binding to DNA by the polymerase (Conaway and Conaway, 1990). *In vitro* studies with purified factors indicates that TFIIF is released from the polymerase complex prior to the release of TFIIH. TFIIF reassociates with a stalled polymerase complex to promote elongation, but appears to be released again when the complex resumes elongation (Price et al., 1989; Kephart et al., 1992; Zawel et al., 1995). In its role as an elongation factor, TFIIF has been demonstrated to increase the elongation rate of RNA polymerase II (Bengal et al., 1991; Kephart et al., 1992; Tan et al., 1995) and to suppress pausing by the polymerase (Price et al., 1989; Tan et al., 1995). Mutagenesis and deletion analysis are defining the regions of the two subunits required for the various functions of TFIIF and suggest that the smaller RAP30 subunit is required for initiation while the larger RAP74 subunit is required for early elongation (Tan et al., 1995; Wang and Burton, 1995; Chang et al., 1995).

The three-subunit transcription elongation factor Elongin (SIII) has been purified from rat tissue (Conaway et al., 1993; Bradsher et al., 1993a, 1993b), and the genes for each subunit cloned from mammalian sources (Garrett et al., 1994a, 1994b; Aso et al., 1995). Like TFIIF, this factor stimulates an increase in the elongation rate of RNA polymerase II. *In vitro*, the Elongin A subunit is absolutely required for this effect and alone is sufficient for partial stimulation. The Elongin C subunit binds directly to A and increases its stimulatory activity. The Elongin B subunit binds only to the C subunit and is a member of the ubibiquitin homology family (Garrett et al., 1995). Elongin B promotes assembly and thermostability of the Elongin complex (Aso et al., 1995). It has been proposed that Elongin functions to maintain correct positioning of the 3'-hydroxyl terminus of the nascent

RNA in the active site of RNA polymerase II and thus suppresses transient pausing by the polymerase (Takagi et al., 1995).

Recently, it has been discovered that the von Hippel-Lindau (VHL) gene, which is mutated in a rare disorder that predisposes individuals to a variety of cancers, encodes a tumor suppressor protein that binds directly to the Elongin B and C subunits. When the B-C subcomplex is bound by VHL protein *in vitro*, it cannot bind to and stimulate the activity of the Elongin A subunit (Duan et al., 1995); the interaction between VHL and the B-C subcomplex has also been demonstrated *in vivo* (Kibel et al., 1995). The B and C subunits bind to a region of VHL that is frequently mutated in human tumors and VHL proteins mutated in this region do not bind to the B-C subcomplex. Thus the tumor suppression activity of VHL is strongly linked to its ability to bind to Elongin B and C (Krumm and Groudine, 1995). The *in vivo* targets of regulation by Elongin are not known. However, as described above, many genes which control cellular proliferation are regulated at the level of transcriptional elongation.

Early Biochemistry of TFIIS Proteins (mouse S-II, yeast P37)

TFIIS was the first eukaryotic transcription elongation factor identified. It was originally identified in mouse as the S-II protein (Sekimizu et al., 1976) and was independently identified in *S. cerevisiae* as the P37 protein (Sawadogo et al., 1979). A brief overview of the early biochemistry of these two factors is relevant. The initial goal of the work described here was to determine if the P37 protein belonged in the TFIIS family of transcription elongation factors.

The transcription elongation factor S-II was originally identified in mouse Ehrlich ascites tumor cells as an activity that stimulated RNA polymerase II in a

nonspecific transcription assay (Sekimizu et al., 1976). The stimulatory activity was purified to a nearly homogeneous preparation of a basic 38 kDa polypeptide and was reported to stimulate both initiation and elongation in nonspecific assays (Sekimizu et al., 1976). The reported stimulation of initiation was probably an artifact of the assay conditions. An antibody generated against S-II inhibited its ability to stimulate RNA polymerase II *in vitro* (Sekimizu et al., 1979b) and also inhibited α-amanitin sensitive transcription in nuclei isolated from Ehrlich ascites tumor cells (Ueno et al., 1979). Immunofluorescence with this antibody indicated that the protein is found in the nucleoplasm but not in the nucleolus or cytoplasm of several normal cell types (Sekimizu et al., 1979a).

Another peak of activity from the S-II purification was named S-II' (Sekimizu et al., 1976). Antibodies against S-II also recognize S-II' (Sekimizu et al., 1979b). Treatment of S-II' with alkaline phosphatase converts it to a polypeptide with a mobility identical to that of S-II suggesting that S-II' is a phosphorylated form of S-II, primarily on serine residues as determined by thin layer chromatography (Sekimizu et al., 1981). *In vitro*, casein kinase II can phosphorylate S-II (Horikoshi et al., 1987). The relevance of this remains unclear since the addition of casein kinase II protects RNA polymerase II from the inhibition of transcription caused by DRB (Zandomeni et al., 1986; Horikoshi et al., 1987), but the addition of purified S-II' does not (Hirai et al., 1988). However, finding that S-II is present in both phosphorylated and unphosphorylated forms *in vivo* suggests that its activity may be regulated by phosphorylation.

The interaction of the TFIIS proteins with RNA polymerase II remains difficult to measure due to the weakness or brevity of the interaction. However, the early characterization of S-II demonstrated that in the absence of DNA, purified

S-II will bind to purified RNA polymerase II in low salt concentrations although the complex is not completely stable to glycerol gradient centrifugation. Pure S-II does not bind to DNA, yet it binds to RNA polymerase II more strongly in the presence of DNA (Horikoshi et al., 1984). Chymotrypsin digestion of S-II in the presence of DNA, or in the presence of RNA polymerase II and DNA, divides the protein into two fragments (Sekimizu et al., 1984). The 18 kDa amino-terminal fragment has no effect on transcription but contains the phosphorylation sites; the 21 kDa carboxyl-terminal fragment retains the ability to stimulate transcription (Horikoshi et al., 1985).

Factors similar to the mouse S-II protein have been purified from yeast (described below), Drosophila (Sluder et al., 1989), and other eukaryotic tissues, including bovine (Rappaport et al., 1987), human (Reines et al., 1989), and rat (Reines, 1992). These proteins are similar in size and in biochemical properties. Comparison of the predicted amino acid sequences of the genes from human (Chen et al., 1992; Yoo et al., 1991), mouse (Hirashima et al., 1988), Drosophila (Marshall et al., 1990) and yeast (Hubert et al., 1983; Ruet et al., 1990; Nakanishi et al., 1992) indicated significant conservation of amino acid sequence, especially in the carboxyl-terminus.

These elongation factors are now considered to be members of the TFIIS family. The TFIIS proteins stimulate the activity of RNA polymerase II in nonspecific transcription assays. Further characterization distinguished the activity of TFIIS from that of TFIIF. TFIIS does not stimulate the elongation rate of RNA polymerase II but promotes an "antitermination" activity (Bengal et al., 1991). In a transcription assay using a specific DNA template containing an intrinsic pause site recognized by purified human RNA polymerase II (Dedrick et al., 1987; Kerppola

and Kane, 1983, 1990), the addition of purified TFIIS stimulates RNA polymerase II to continue to elongate through the pause site (Reinberg and Roeder, 1987; Rappaport et al., 1987; Reines et al., 1989; SivaRaman et al, 1990; Sluder et al., 1989). A monoclonal antibody directed against a region (homology block B) of the largest subunit of RNA polymerase II prevents the stimulatory activity of TFIIS, suggesting that this elongation factor interacts with this highly conserved region of the polymerase (Rappaport et al., 1988).

A similar stimulatory activity called P37 was identified in *S. cerevisiae* as a protein which stimulated the activity of RNA polymerase II in a nonspecific transcription assay (Sawadogo et al., 1979). Careful analysis of the activity of P37 indicated that it stimulates the elongation reaction of RNA polymerase II and does not stimulate the initiation reaction (Sawadogo et al., 1981). P37 co-sediments through a glycerol gradient with yeast RNA polymerase II and was reported to bind preferentially to RNA polymerase II containing an intact carboxyl-terminal domain (CTD) of the largest subunit over polymerase containing a proteolyzed CTD. This result suggested that P37 might interact with the CTD (Sawadogo et al., 1979) .

However, other RNA polymerase subunits may also be necessary for the interaction of TFIIS with the polymerase. Antibodies generated against the RPB6 subunit, common to all three yeast nuclear RNA polymerases, inhibit transcription by RNA polymerase II. Preincubation of RNA polymerase II with the P37 protein prior to addition of the antibody protects against this transcriptional inhibition (Sawadogo et al., 1980) and suggests that the elongation factor interacts with the RPB6 subunit. Based on the biochemical similarity of yeast P37 to the TFIIS proteins in nonspecific transcription assays, it seemed likely that the P37 protein

might be a member of the TFIIS family. Biochemical demonstration of this functional homology and a preliminary investigation into the polymerase contacts made by the yeast TFIIS protein (P37) is described in Chapter 2 of this work.

Cleavage of the Nascent RNA Transcript

Among the small group of proteins demonstrated to stimulate elongation by RNA polymerase II, the TFIIS proteins possess a unique function. Unlike TFIIF and Elongin, TFIIS does not stimulate the overall rate of elongation by the polymerase (Bengal et al., 1991). Unlike TFIIH which promotes a transition from an elongation incompetent form to an elongation competent one, TFIIS does need to be present at the time of initiation to exert its effect (SivaRaman et al., 1990). The TFIIS proteins stimulate stalled elongation complexes of RNA polymerase II to continue elongating and part of the mechanism involves inducing cleavage of the nascent RNA transcript (reviewed in Kassavetis and Geiduschek, 1993; Reines, 1994).

TFIIS stimulates RNA polymerase II stalled by a variety of obstacles to elongation to cleave the nascent transcript. At some types of blocks to elongation, such as DNA sequences that possess an intrinsic ability to cause the polymerase to pause (SivaRaman et al., 1990; Reines et al., 1989; Reines et al., 1992; Wang and Hawley, 1993; Izban and Luse, 1993b) or DNA-binding proteins that physically stop the polymerase (Reines and Mote, 1993), the addition of TFIIS stimulates cleavage of the nascent transcript, followed by read-through of the block. The addition of TFIIS to stalled ternary elongation complexes containing RNA polymerase II halted by other methods, such as nucleotide limitation (Izban and Luse, 1992, 1993a, 1993b) or a DNA lesion (Donahue et al., 1994), also stimulates

transcript cleavage. In these cases, TFIIS is not sufficient to promote continued elongation, which requires readdition of nucleotides or removal of the lesion.

The TFIIS-dependent cleavage of the transcript is thought to be an essential step in the mechanism by which TFIIS promotes read-through of blocks to elongation (Reines, 1992, 1994). However, the cleavage reaction itself is carried out by the polymerase (Rudd et al., 1994; R. Weilbaecher, unpublished results), and cleavage is not sufficient for read-through (Ciprés-Palacín and Kane, 1994, 1995). The discovery of the TFIIS-dependent transcript cleavage reaction of RNA polymerase II closely followed the discovery of transcript cleavage by *E. coli* RNA polymerase (Surratt et al., 1991). In *E. coli*, there are two factors which promote transcript cleavage, GreA and GreB (Borukhov et al., 1992, 1993; Feng et al., 1994). Investigation into the transcript cleavage reaction of RNA polymerases has paralleled the development of a new model for the mechanism of polymerization of RNA chains and movement along the template by RNA polymerase (Chamberlin, 1994).

A previous model (von Hippel and Yager, 1992; Yager and von Hippel, 1987) viewed the polymerase as a relatively static macromolecule proceeding along the DNA template in monotonic manner. For each phosphodiester bond formed, the polymerase opened one DNA:DNA base pair on the leading edge and closed one on the lagging edge of the opened region of DNA. In this region, eighteen base pairs of the template were opened in the transcription "bubble". Approximately twelve nucleotides at the 3'-growing end of the nascent RNA transcript were base paired with the template DNA strand. This RNA:DNA hybrid was considered to be required for the stability of the ternary elongation complex; disruption of the hybrid was viewed as a cause of transcription termination.

However, this model was inadequate to explain the release of the 3'-growing end of the nascent RNA in the transcript cleavage reaction of RNA polymerases. It also could not explain a large amount of other experimental information about elongation and termination (reviewed in Chamberlin, 1994).

A new model of the mechanism of RNA polymerases views the polymerase as a dynamic structure that moves along the template DNA like a molecular inchworm. RNA polymerases, unlike DNA polymerases, are processive enzymes. Since RNA polymerase cannot reassociate with a released RNA transcript and continue its polymerization, it is imperative that the polymerase does not let go of the template or the nascent RNA. Thus the new model proposes that the polymerase contains two DNA-binding sites and two RNA-binding sites. One of each type of site is always tightly associated as the polymerase moves along the template. Another key feature of this model is that the catalytic site of the enzyme moves with respect to the position of the RNA-binding sites. This model explains certain types of stalled complexes as being unable to continue elongation due to a misalignment of the catalytic site and the 3'-growing end of the nascent RNA. Cleavage of the nascent RNA by the catalytic site generates a new 3'-growing end of the RNA that is correctly aligned with the catalytic site and with the template DNA (Chamberlin, 1994; Nudler et al., 1994).

The ability of RNA polymerases to cleave the nascent transcript appears to be an intrinsic property of RNA polymerases. To date, every multisubunit RNA polymerase examined possesses this ability. The cleavage reaction appears to be catalyzed by the active site of the polymerase (Rudd et al., 1994). Both purified *E. coli* RNA polymerase and purified *S. cerevisiae* RNA polymerase II can cleave the nascent RNA in the absence of accessory factors under some conditions (Orlova et

al., 1995; D. Solow-Cordero unpublished results; R. Weilbaecher, unpublished results). However, the cleavage reaction is greatly stimulated by the addition of the cleavage factors, GreA and GreB for *E. coli* RNA polymerase or TFIIS for eukaryotic RNA polymerase II. It is proposed that TFIIS promotes a conformational change in the polymerase that is required for transcript cleavage (Reines, 1994). TFIIS may also be required to stabilize the conformation of the polymerase to promote read-through of blocks to elongation (Ciprés-Palacín and Kane, 1994, 1995)

The Structure of TFIIS

Currently the interaction of TFIIS with RNA polymerase II is not well understood although significant progress has been made. For every TFIIS protein characterized, the carboxyl-terminal portion is sufficient for the known biochemical activities, stimulation of transcript cleavage and read-through of blocks to elongation (Horikoshi et al., 1985; Agarwal et al., 1991; Reines et al., 1992; Guo and Price, 1993; Ciprés-Palacín and Kane 1994; Christie et al., 1994; Nakanishi et al., 1995; Chapters 2 and 3 of this report). This portion of the TFIIS protein appears to be composed of two structural elements, a series of α helices (D. Awrey and A. Edwards, personal communication) and a zinc-ribbon domain (Qian et al., 1993a, 1993b). It has been proposed that the zinc-ribbon domain may be part of a nucleic acid binding domain that is normally masked by the amino-terminus of the TFIIS protein and that the middle region of the TFIIS protein may contain an RNA polymerase binding domain (Agarwal et al., 1991). However, there is little evidence for either of these functional assignments. Another possibility is that the TFIIS protein exerts its effect solely through modulation of the

conformation of RNA polymerase. Further investigation into the interaction of TFIIS with the polymerase will distinguish between these possibilities.

RNA polymerase II

In order to understand how TFIIS exerts its effect on polymerase, it is necessary to understand the polymerase itself. While the general initiation factors have been studied extensively to assign individual functions to each factor and sometimes even each subunit, far less is known about the functions of the subunits of RNA polymerase II. Eukaryotic RNA polymerases are more complex enzymes than RNA polymerase from *E. coli*; yeast RNA polymerase II contains twelve subunits (Young, 1991; Woychik and Young, 1990; Thuriaux and Sentenac, 1992). Each has been cloned, but unlike the case with *E. coli* RNA polymerase (Borukhov and Goldfarb, 1993) it is not currently possible to reconstitute active eukaryotic enzyme from individually expressed subunits. Thus it has been difficult to elucidate functions for the small subunits of eukaryotic RNA polymerases.

Comparison of the cloned sequences of the two largest RNA polymerase subunits from eukaryotes, archaebacteria, and poxviruses with the β and β' subunits of *E. coli* RNA polymerase has revealed significant regions of homology that are conserved in all multisubunit RNA polymerases (Allison et al., 1985; Broyles et al., 1986; Sweetser et al., 1987; Puhler et al., 1989; reviewed in Thuriaux and Sentenac, 1992). These similarities suggest that the two largest subunits may form a core containing DNA-binding, RNA-binding, nucleotide-binding, and zinc-binding activities (Thuriaux and Sentenac, 1992). Mutants in *S. cerevisiae* are beginning to confirm the involvement of the two large subunits in these functions. A mutation in the largest subunit of RNA polymerase III causes a defect in

elongation and implicates a motif invariant amongst the multisubunit RNA polymerases in formation of the active site (Dieci et al., 1995). Mutations in the second largest subunit of RNA polymerase III cause defects in termination (Shaaban et al., 1995).

The third largest subunit of RNA polymerase II, RPB3, may also form a part of the "core enzyme". RPB3, like the *E. coli* α subunit, is present in a stoichiometry of two polypeptides per enzyme and appears to be close to the active site. Certain mutations in each of the three largest subunits of yeast RNA polymerase II, RPB1, RPB2, or RPB3, cause defects in enzyme assembly (Kolodziej and Young, 1991). Although the homology of RPB3 to the *E. coli* α subunit is limited, mutations in this region of the *E. coli* polypeptide block subunit assembly (Igarashi et al., 1990). The RPB11 subunit also contains homology to this region of the *E. coli* α subunit (Woychik et al., 1993). It is proposed that this region may be involved in interactions of the subunits of RNA polymerase II (Woychik and Young, 1994).

The largest subunit of eukaryotic RNA polymerase II enzymes also contains a carboxyl-terminal domain (CTD) composed of multiple heptapeptide repeats with a consensus sequence of YSPTSPS (Allison et al., 1988; Corden, 1990). In yeast, partial deletions of this domain confer defects in the transcription of certain promoters and complete deletions are lethal (Nonet et al., 1987; Nonet and Young, 1989; Scafe et al., 1990). Phosphorylation of this domain, also essential in yeast (Valay et al., 1995; West and Corden, 1995), converts the polymerase from the hypophosphorylated IIa form to the hyperphosphorylated IIo form. The IIa form is recruited to the promoter by the general initiation factors (Laybourn and Dahmus, 1990; Dahmus, 1994). Phosphorylation of the CTD by TFIIH (described above) is implicated in the regulation of promoter escape/clearance. The CTD also interacts

directly with other general initiation factors, including TFIID (Usheva et al., 1992; Thompson et al., 1993). The CTD also plays a role in the formation of RNA polymerase II holoenzyme complexes (Koleske and Young, 1994, 1995; Kim et al., 1994a). The large holoenzyme complex contains the twelve subunits of RNA polymerase II and many other proteins which mediate the ability of the polymerase to respond to transcriptional activators. The activities of this complex during *in vivo* regulation of transcriptional initiation and elongation are not yet clear.

Homologies with the well characterized *E. coli* enzyme cannot be used to suggest functions for the remaining subunits, but a variety of mutants in *S. cerevisiae* have begun to elucidate the roles of these small subunits. Two small subunits, RPB4 and RPB7, form a subcomplex and are present in substoichiometric amounts in purified preparations of RNA polymerase II (Kolodziej et al., 1990). Deletion of RPB7 is lethal (McKune et al., 1993); deletion of RPB4 is not lethal but causes temperature sensitivity and defects in stress responses (Woychik and Young, 1989). RNA polymerase II purified from a strain deleted for RPB4 does not contain RPB7 either and is defective in specific initiation but not in elongation *in vitro* (Edwards et al., 1991). This subcomplex may render the polymerase more resistant to environmental stress since the proportion of RNA polymerase II enzymes containing these two subunits increases as cells enter stationary phase (Choder, 1993; Choder and Young, 1993). The human homologue of RPB7 rescues the lethality of an RPB7 null at moderate temperatures, but not at temperature extremes or during maintenance in stationary phase (Khazak et al., 1995).

Like the RPB4 subunit, the RPB9 subunit is required for normal growth at extremes of temperature (Woychik et al., 1991). Deletion the RPB9 subunit also

causes defects in start site selection and the zinc-binding domain of this subunit appears to play a role in this function (Furter-Graves et al., 1994; Hull et al., 1995). These defects are distinct from the defects conferred by mutation of TFIIB (Pinto et al., 1992) or of RPB1 (Hekmatpanah and Young, 1991). The RPB9 zinc-binding domain shares significant homology with the zinc-ribbon domain of TFIIS (Qian et al., 1993a), although structural analysis has not yet confirmed that its structure is similar. Nevertheless, in combination with the start site defect of some RPB9 mutants, this similarity suggests the possibility that this zinc-binding domain may play a role in positioning the catalytic site of the polymerase. Indeed some of the elongation properties of RNA polymerase II lacking this subunit are different than those of the wild type enzyme (R. Weilbaecher, unpublished results).

In *S. cerevisiae*, there are five subunits that are common to all three nuclear RNA polymerases, encoded by the essential, single-copy genes *RPB5, RPB6, RPB8, RPC10,* and *RPB10* (Woychik et al., 1990; Carles et al., 1991). At least two of these subunits, RPB6 and RPB10 are highly conserved amongst eukaryotes and archaebacteria (McKune and Woychik, 1994a). The mammalian homologues of *RPB6, RPB8, RPC10,* and *RPB10* are functional in *S. cerevisiae* (McKune and Woychik, 1994b; Shpakovski et al., 1995). The RPB6 subunit is of particular interest in the study of TFIIS function and its effect on transcriptional elongation.

Interactions Between TFIIS and RNA polymerase II

It is clear that the TFIIS elongation factor interacts directly with RNA polymerase II. *In vitro* it is capable of interacting with purified polymerase in the absence of any accessory factors (SivaRaman et al., 1990; Guo and Price, 1993; Christie et al., 1994; Chapters 2 and 3 of this work). However, the specific contacts

between this elongation factor and the polymerase are not known. The early characterization of TFIIS proteins (described above) provided biochemical evidence that suggested that TFIIS interacts with the largest subunit and the common subunit RPB6 (Sawadogo et al., 1980; Rappaport et al., 1988). More recently, genetic evidence has implicated these same two subunits as interacting with TFIIS. Deletion of the *PPR2* gene encoding the *S. cerevisiae* TFIIS protein results in sensitivity to the uracil analogue 6-azauracil (6-AU; Hubert et al., 1983; Exinger and Lacroute, 1992). Certain mutations in *RPB1* (Archambault et al., 1992b) also cause sensitivity to 6-AU. These *RPB1* mutations are clustered in a region of the subunit (homology block H) conserved among eukaryotic RNA polymerases and most of them also confer a temperature sensitive phenotype. The 6-AU sensitivity conferred by these RPB1 mutations can be suppressed by overexpression of the *PPR2* gene (Archambault et al., 1992b). Other mutations in this region of RPB1 can be suppressed by overexpression of the RPB6 subunit. The allele specificity of this suppression suggests a direct interaction between the RPB1 and RPB6 subunits (Archambault et al., 1992a; Archambault et al., 1990). The high degree of conservation of these three polypeptides, encoded by *RPB1*, *RPB6*, and *PPR2*, and the effects of mutations in these genes on transcriptional elongation implicates these proteins in an important role in the polymerization of RNA. Elucidation of the interaction of TFIIS with these two subunits of RNA polymerase II will further our understanding of the mechanism of RNA polymerases.

TABLE LEGEND

Table 1 - Genes Reported to Contain Blocks to Transcriptional Elongation

Genes reported to contain blocks to transcriptional elongation are listed along with references. The list is divided into three sections according to the reported location of the block to elongation. The c-*myc* and c-*fos* genes are listed in two sections. As explained in the text in the case of the c-*myc* gene, further examination of the genes in section B may be warranted. **A** - The genes listed in this section have been reported to contain promoter proximally paused polymerases very early within the transcription unit. Generally these pauses are 20-40 base pairs downstream of the initiation site. **B** - The genes listed in this section have been reported to contain blocks to elongation within the body of the gene. Most of these have been localized to sequences in the exon 1/intron 1 junction. A few have been localized to sequences within the second intron. **C** - These genes have been reported to contain blocks to elongation in the 3'-region of the gene. Most of these are downstream of the poly(A) site and are now thought to play a role in proper 3'-end formation.

A - Promoter Proximal Blocks (20-40 bases downstream of start site)

Drosophila hsp70	Gilmour and Lis, 1985; Gilmour and Lis, 1986; Rougvie and Lis, 1988
Drosophila hsp26	Rougvie and Lis, 1990
β-1-tubulin	Rougvie and Lis, 1990
polyubiquitin	Rougvie and Lis, 1990
glyceraldehyde-3-phosphate dehydrogenase-1	Rougvie and Lis, 1990
glyceraldehyde-3-phosphate dehydrogenase-2	Rougvie and Lis, 1990
copia	referred to in Rougvie and Lis, 1988
α-tubulin	Middleton and Morgan, 1990; Hair and Morgan, 1993
transthyretin	Mirkovitch and Darnell, 1992
c-*myc*	Wright et al., 1991; Strobl and Eick, 1992; Krumm et al., 1992; Strobl et al., 1993; Wolf et al., 1995
c-*fos*	Plet et al., 1995

B - Intragenic Blocks to Elongation, Usually Within the First Intron

c-*myc* Bentley and Groudine, 1986; Eick and Bornkamm, 1986; Nepveu and Marcu, 1986

L-*myc* Krystal et al., 1988; Xu et al., 1991

N-*myc* Krystal et al., 1988; Xu et al., 1991

c-*myb* Bender et al., 1987; Watson, 1988

c-*fos* Fort et al., 1987; Mechti et al., 1991

c-*fms* Yue et al., 1993

EGF receptor (c-*erbB*)	Haley and Waterfield, 1991
Adenosine deaminase	Chinsky et al., 1989; Chen et al., 1990; Kash and Kellems, 1994
Porphobilinogen deaminase	Beaumont et al., 1989; Porcher et al., 1995
Ribonucleotide reductase	Björklund et al., 1992
Ornithine decarboxylase	Shor et al., 1995
histone H3.3 (human)	Reines et al., 1989; D. Wells, unpublished results
Chicken β^A globin	Lois et al., 1990
Retinoic Acid Receptor Isoforms β1 and β3	Mendelsohn et al., 1994
Androgen receptor	Grossmann and Tindall, 1995
Opsin (bovine)	DesJardin et al., 1995
Tumor Necrosis Factor α (TNF-α)	Biragyn and Nedospasov, 1995
Cystic Fibrosis Transmembrane Regulator (CFTR)	Koeberl et al., 1995
Clotting Factor VIII	Koeberl et al., 1995

C - Genes Containing Blocks/Termination Sites in the 3'-Region

Immunoglobulin κ	Xu et al., 1986
Immunoglobulin heavy chain μ-δ spacer	Mather et al., 1984
Immunoglobulin μ	Galli et al., 1987
β-globin	Gariglio et al., 1981 Villeponteau et al., 1982
Mouse	Hofer and Darnell, 1981; Hofer et al., 1982; Citron et al., 1984; Falck-Pederson et al. 1985
Chicken	Groudine et al., 1981
Rabbit	Rohrbaugh et al., 1985
α-globin	
Chicken	Weintraub et al., 1981
Mouse	Shefferey et al., 1984
Human	Whitelaw and Proudfoot, 1986
Rabbit	Vandenbergh et al., 1991
α-amylase (mouse)	Hagenbüchle et al., 1984
Ovalbumin	LeMeur et al., 1984
Dihydrofolate reductase	Frayne et al., 1984
(mouse)	Frayne et al., 1986
Calcitonin	Amara et al., 1984
Gastrin (human)	Sato et al., 1986; Baek et al., 1986
histones	Birchmeier et al., 1984
sea urchin H2A	Johnson et al., 1986
U1 and U2 snRNAs	Hernandez and Wiener, 1986 Neuman de Vegvar et al., 1986 Hernandez and Lucito, 1988 Parry et al, 1989
human complement C	Ashfield et al., 1991; Ashfield et al., 1994; Moreira et al., 1995
S. pombe ura4 gene	Humphrey et al., 1994
Trypanosome splice leader	Kooter et al., 1984; Campbell et al., 1984

Chapter 2

Comparison of Full Length and Truncated Yeast TFIIS with Several Variants of RNA polymerase II *in vitro*

This chapter has been previously published in slightly modified form as:

Christie, K.R., Awrey, D.E., 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, P37. J. Biol. Chem. **269**, 936-943.

INTRODUCTION

Many eukaryotic genes contain regulated blocks to transcript elongation within the transcription unit (Kerppola and Kane, 1991; Spencer and Groudine, 1990). 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 (Kerppola and Kane, 1991; Bengal et al., 1991; Reinberg and Roeder, 1987; Reines et al., 1989; SivaRaman et al., 1990; Sluder et al., 1989; Wiest et al., 1992).

TFIIS-related proteins which stimulate elongation by RNA polymerase II have been purified and the genes cloned from a variety of organisms (Sawadogo et al., 1979; Hubert et al., 1983; Ruet et al., 1990; Nakanishi et al., 1992; Sluder et al., 1989; Marshall et al., 1990; Rappaport et al., 1987; Chen et al., 1992; Hirashima et al., 1988; Kanai et al., 1991; Yoo et al., 1991). Sequence analysis of these clones predicts significant amino acid similarities especially in the C-terminal region of the proteins (Chen et al., 1992). In most cases, these TFIIS-related proteins have also been shown to promote read-through of specific blocks to elongation *in vitro* (Bengal et al., 1991; Reinberg and Roeder, 1987; Reines et al., 1989; SivaRaman et al., 1990; Sluder et al., 1989; Wiest et al., 1992).

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 (Sawadogo et al., 1979; Reinberg and Roeder, 1987; Sluder et al., 1989; Agarwal et

al., 1991; Horikoshi et al., 1984), although some data suggest that they do not remain associated throughout elongation (Sluder et al., 1989). Mutants in mouse SII (Horikoshi et al., 1990) and human TFIIS (Agarwal et al., 1991) 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 (Reines, 1992; Reines et al., 1992; Izban and Luse, 1992; Wang and Hawley, 1993; Kassavetis and Geiduschek, 1993; Guo and Price, 1993). 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 (Guo and Price, 1993; SivaRaman et al., 1990; Yoo et al., 1991; Agarwal et al., 1991; Natori, 1982), but catalytically active RNA polymerase II purifies as a complex of nine to twelve polypeptides (Sentenac, 1985; Kolodziej et al., 1990).

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 (Rappaport

et al., 1988). 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 (Archambault et al., 1992b). Also, the yeast TFIIS analogue apparently binds more efficiently to RNA polymerase II when its largest subunit contains the highly conserved C-terminal domain (CTD) of heptapeptide repeats (Sawadogo et al., 1979). This domain, unique to RNA polymerase II (Allison et al., 1988; Corden, 1990), is often removed by proteolysis during purification to generate form IIb (Sentenac, 1985). In this work, we examine the TFIIS-stimulated read-through activity of purified RNA polymerase II with or without the conserved C-terminal domain of the largest subunit.

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 (Sawadogo et al., 1980). 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 (McKune et al., 1993). Purified yeast RNA polymerase II contains substoichiometric amounts of these two subunits (Woychik and Young, 1990; Edwards et al., 1991). While 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* (Edwards et al., 1991). 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 (Reines et al., 1987), and mammalian TFIIS stimulates mammalian polymerase to read-through this block (Reines et al., 1989; SivaRaman

et al., 1990). 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.

MATERIALS AND METHODS

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. Nucleotides were obtained from Pharmacia. [α-^{32}P]-CTP, >400 μCi/mmole was obtained from Amersham. BioGel 30 columns were obtained from BioRad.

RNA polymerase II forms IIa and Δ4,7 were the generous gift of A. Edwards, McMaster University, Ontario, Canada and were purified by immunoaffinity as described (Thompson et al., 1990; Edwards et al., 1990). 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 (Li and Kornberg, 1994). The specific activities of the IIa and Δ4,7 polymerases were comparable. The specific activity of the ΔCTD form was approximately two to four 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[1] *Saccharomyces cerevisiae* genomic BamHI fragment containing the *PPR2* gene, encoding TFIIS, from pFL44D (generously provided by A. Guyonvarch, A. Ruet, and F. LaCroute) was ligated into the BamHI site of pBS.KS$^+$ (Stratagene) to generate pKC3. Plasmids were then constructed encoding fusion proteins between an amino-terminal histidine-rich region and 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 NdeI and BamHI, and the fragment was ligated into the 5.7 kb NdeI-BamHI digested pET15b (Novagen). This plasmid is referred to as pET15bPPR2.

The truncated reading frame PCR product was digested with BamHI, and the fragment was ligated into the 3.0 kb BamHI-SmaI digested pGEM7Zf(+) (Promega). Next, the 0.6 kb NdeI-BamHI 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 *E. coli* BL21 (DE3) cells which were then grown to an OD_{600} of 0.8 to 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 three hours before harvesting and were resuspended in 50 mM Tris-HCl, pH 7.5 and 10% sucrose (2 ml per gram 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 phenylmethylsulfonylfluoride, 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 minutes at 4°C. The suspension was centrifuged in a Sorvall GSA rotor at 11,000 rpm for 20 minutes at 4°C. The supernatant solution was loaded onto a 5 x 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 x 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[1], 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 minutes 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 M to 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 mM to 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 MonoS 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} of 0.8 to 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 BioRad #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 x 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 M to 1.0 M potassium acetate.). Throughout the purification, the presence of full length or truncated TFIIS was monitored using SDS[1]/poly-acrylamide gel electrophoresis.

Preparation from *Saccharomyces cerevisiae* of TFIIS/P37-Containing Fraction

A partially purified fraction of TFIIS/P37 protein was obtained essentially as described (Sawadogo et al., 1979). Briefly, a clarified cell lysate was prepared from 80 g of *Saccharomyces 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 by Sawadogo and Sentenac (Sawadogo et al., 1979). 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 (Sawadogo et al., 1979).

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 (E. Rosenberg and H. Nikaido, unpublished procedure).

Silver Staining of SDS/Polyacrylamide Gels

Proteins were visualized by the silver staining protocol of E. Rosenberg and H. Nikaido. SDS/polyacrylamide gels were fixed by shaking 1 hr in 30% ethanol/10% acetic acid. Gels were then incubated for one hour in 0.5% glutaraldehyde, 0.1% sodium thiosulfate, 30% ethanol, and 0.4 M NaOAc, pH 6.0. Excess glutaraldehyde was removed by soaking the gel in distilled, deionized water overnight. Gels were then rocked in 0.1% $AgNO_3$/0.01% fomaldehyde for 45'. The gel was rinsed once briefly in developer and then soaked in developer

until bands appeared, generally 3' or less. Developer was poured off and the gel was rocked in 5% acetic acid for 15', two times, to prevent further development. Acetic acid was removed by soaking the gel in H_2O. The gel was dried between sheets of cellulose acetate (BioRad).

Preparation of 3'-Deocycytidine-Extended Template pCpGEMTerm

The pGEMTerm plasmid contains a 300 base pair TaqI fragment from intron 1 of the human histone H3.3 gene containing three intrinsic blocks to elongation (Reines, 1987 and see Figure 2). DNA for the preparation of 3'-deocycytidine-extended templates were prepared by alkaline lysis (as described in Chapter 3, except for the use of a 1 L culture with all subsequent volumes scaled accordingly). The DNA pellet was dissolved in TE (10 mM Tris-Cl, pH 8, 1 mM EDTA) and the final volume determined. The plasmid DNA was purified by two sequential equilibrium bandings through cesium chloride gradients (1.09 g cesium chloride/ml volume and 1/10 volume 2 mg/ml ethidium bromide were added to the DNA solution).

The pGEMTerm plasmid DNA (100 µg) was linearized by digestion with the enzyme SmaI (NEB) until electrophoresis of a sample indicated complete cutting. The restriction enzyme was removed by phenol extraction. The DNA was precipitated with NH_4OAc (0.8M) and ethanol, washed with 70% ethanol, and resuspended in 105 µl TE; a 5 µl sample was removed and the remainder was prepared as template.

The 3'-ends of the DNAs were extended with terminal transferase (Ratliff) in standard conditions for addition of deoxycytidine (200 mM potassium cacodylate, pH 7.5; 1 mM CoCl2; 2 mM βME; and 20 µM deoxycytidine (dCTP); 2 µl 120

Unit/ml terminal transferase) (Kadesch and Chamberlin, 1982). From this point forward, all plastic tubes and tips were coated with silicon (SigmaCote from Sigma) to minimize sticking of the tailed DNA to plastic. A 20 µl aliquot of this reaction was removed and to it was added 0.5 µl [α^{32}P]-dCTP (to a final concentration of approximately 200-1000 cpm/picomole dCTP). The tailing reactions were incubated at 37°C. Aliquots (5 µl) of the labelled reaction were removed at various times to determine the incorporation of deoxycytidine into precipitable counts by precipitation with trichloroacetic acid. For efficient transcription, 3'-extensions of at least 10 residues are required. The tailing reaction was continued, with addition of additional terminal transferase if needed, until the 3'-extensions were sufficient as judged by incorporation of [α^{32}P]-dCMP and by electrophoresis of a sample of the tailed DNA adjacent to the untailed aliquot. The DNA was purified by phenol extraction and ethanol precipitation as described above and dissolved in 50 µl TE.

The final step in template preparation was to cut the 3'-extended DNA to generate a molecule of pairs a convenient size containing only one tail. EcoRI was used to remove the tail from the end not desired for transcription. PstI cuts in the polylinker on the other side of the 300 base pair TaqI fragment and limits run-off transcription to this size. The DNA was purified by phenol extraction and ethanol precipitation as described above and dissolved in 300 µl TE.

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 (Kadesch and Chamberlin, 1982) of the TaqI fragment

containing the human histone H3.3 gene (Reines et al., 1987). In this template are three characterized blocks to elongation, TII, TIb, and TIa (Reines et al., 1987; Kerppola and Kane, 1990). Reactions (50 µl) in transcription buffer [60 mM Tris-OAc, pH8; 5 mM MgOAc; 5% glycerol; 100 mM NH_4OAc; 6 mM spermidine-HCl; 0.8 mM each ATP, GTP, and UTP; 1 µM [α^{32}P]-CTP; 1U Inhibit-ACE; and yeast RNA polymerase II] were incubated for 1 minute at 30°C. A 5 µl aliquot was diluted into 100 µl stop buffer [20 mM EDTA; 0.5 M NH_4OAc; and 12 µg/ml carrier RNA] on ice for analysis. At 1 minute 15 sec, 30 µl was diluted 10x into chase buffer [transcription buffer plus 100 µg/ml heparin and 0.1 mM CTP] and incubated for 45 seconds 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 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 $(NH_4)_2OAc$, and 15% glycerol). Aliquots (30 µl) were removed to 100 µl stop buffer at various timepoints, and RNAs were ethanol precipitated, resuspended in formamide load buffer [90 mM Tris-borate, pH 8.0, 2 mM EDTA, 80% formamide, 0.13% bromophenol blue, 0.13% xylene cyanol, 0.06% SDS] and resolved by electrophoresis through 5% acrylamide-8.3 M urea gels.

Quantitation of Transcripts in Polyacrylamide Gels for Read-through Assays

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 timepoint 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) - (TII cpm) - (TII cpm)].

Ternary Complex Cleavage Assay

Bio-Gel 30 columns were equilibrated by overlaying each column with 200 µl chase buffer lacking nucleotides, and centrifuging for 2 minutes 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 minute at 30°C. The reactions were then diluted 5x into chase buffer at 30°C, and the incubation was continued for one additional minute. 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 minutes 30 seconds 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 a-amanitin was used, it was added to a final concentration of 100 µg/ml for at least one minute 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% bromophenol blue, 0.03% xylene cyanol]. RNA products were resolved by electrophoresis through 5% acrylamide-8.3 M urea gels.

RESULTS

TFIIS Activity in *Saccharomyces cerevisiae*

A partially purified protein from *Saccharomyces cerevisiae* (P37) had been shown to stimulate elongation by both RNA polymerase I and II in non-specific assays (Sawadogo et al., 1979, 1980, 1981). 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 non-specific transcription assays (Sekimizu et al., 1979b). The SII protein is a member of the TFIIS family of proteins, cloned from mouse (Hirashima et al., 1988), human (Chen et al., 1992; Yoo et al., 1991) and *Drosophila* (Marshall et al., 1990). Each of these proteins promotes transcript read-through as well as nascent transcript cleavage (Kassavetis and Geiduschek, 1993). Protein sequences predicted from the gene encoding P37, *PPR2* (Hubert et al., 1983; Ruet et al., 1990; and Nakanishi et al., 1992), place it into the TFIIS family. To see whether this protein was a functional homolog of TFIIS, the previously reported P37 purification (Sawadogo et al., 1979) 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. 1B, Lane 3). This preparation stimulates yeast RNA polymerase II two to five fold in non-specific 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 (Edwards et al., 1991).

Transcripts with 3' ends at TII, TIb, and TIa, are observed forty-five seconds after the addition of chase mixture to initiated complexes (Fig. 2, Lane 2). The TII and TIb sites stop the polymerase with low efficiency, ten percent or less. However, the polymerase recognizes the TIa site with approximately fifty percent efficiency. During continued incubation, some ternary complexes stalled at the site appear to resume elongation, while the majority of the complexes originally stopped at TIa do not read through the site even after thirty minutes of incubation at 30°C (Fig. 2, lane 12). The apparent resumption of elongation in the absence of TFIIS is probably an artifact of the quantitation calculation (see Chapter 3). When TFIIS/P37 protein is added to ternary complexes stalled at these sites, transcript elongation begins within five minutes (Fig. 2, lane 9), and most complexes initially stalled at TIa elongate transcripts to the run-off within thirty minutes (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. 1B, Lane 2) or the truncated TFIIS (Fig. 1B, Lane 1) were used in combination with several variant forms of yeast RNA polymerase II (Δ4,7, IIa, ΔCTD; Fig. 1A). 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 Figure 3; the quantitative analyses for all forms of the polymerase and TFIIS are displayed in Figure 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 ten percent efficiency and the TIa site with about fifty percent efficiency. When either form of TFIIS is added, the complexes resume elongation, and within thirty minutes, eighty percent of the transcripts originally ending at TIa have been elongated. None of the transcripts which remain at TIa after thirty minutes is elongated during continued incubation, and it is likely that these transcripts have been released from the polymerase. In the absence of TFIIS, the majority of the transcripts with 3' ends at TIa have not been elongated, even after fifty minutes of incubation (Figures 3 and 4).

Clearly, neither the C-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 (Agarwal et al., 1991; Ciprés-Palacín and C. Kane, 1994), mouse (Reines et al., 1992; Horikoshi et al., 1985), and *Drosophila* (Guo and Price, 1993) 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 Figure 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 (Reines, 1992; Reines et al., 1992; Izban and Luse, 1992; Wang and Hawley, 1993; Guo and Price, 1993; reviewed in Kassavetis and Geiduschek, 1993). A similar cleavage reaction has been observed in ternary complexes formed with vaccinia virus RNA polymerase (Hagler and Shuman, 1993) which contains a subunit with sequence similarity to the TFIIS family of proteins (Ahn et al., 1990). 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 (Reines, 1992; Reines et al., 1992; Izban and Luse, 1992; Wang and Hawley, 1993; Guo and Price, 1993; reviewed in Kassavetis and Geiduschek, 1993). This cleavage has been observed with complexes stalled by controlling nucleotide levels (Izban and Luse, 1992; Hagler and Shuman, 1993; Izban and Luse, 1993a, Izban and Luse, 1993b), by a DNA binding protein (Reines and Mote, 1993), or by naturally occurring blocks to elongation (Reines, 1992;

Reines et al., 1992; Wang and Hawley, 1993; Izban et al., 1993b). 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 (Reines et al., 1992). 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 five seconds 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 is 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 a-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 (Hagler and Shuman, 1993) or complexes with mammalian or *Drosophila* RNA polymerase II in response to TFIIS (Reines et al., 1992; Izban and Luse, 1992; Wang and Hawley, 1993; Guo and Price, 1993). The exact sizes of the primary cleavage products removed from the 3' end in yeast ternary complexes are not yet known.

Within ten seconds 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 polymerase II ratio of 5:1, most of the transcripts with 3' ends at TIa have been shortened within five minutes (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 ten seconds 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 ten minutes of incubation (Fig. 4B).

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 (Wieland and Faulstich, 1991). 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 (Chamberlin, 1994), and α-amanitin may not interfere with the cleavage reaction although it may block elongation at some other step such as polymerase translocation (Wieland and Faulstich, 1991; Gu et al., 1993). 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 (Edwards et al., 1990) taking advantage of antibody directed against the C-terminal repeated sequence unique to the largest subunit of RNA polymerase II (Sentenac, 1985; Corden, 1990). 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. 1A).

DISCUSSION

The conserved C-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 (Nonet et al., 1987). Further, deletion of subunit four causes slow growth and temperature sensitivity of the cells (Woychik and Young, 1989) while deletion of subunit seven is lethal (McKune et al., 1993). 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 (Reines et al., 1992; Izban and Luse, 1992; Wang and Hawley, 1993; Guo and Price, 1993; and this report). Since this toxin binds to the polymerase, this inhibition suggests that transcript cleavage might be catalyzed by the polymerase 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 ten seconds 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 fifty percent of RNA polymerase II molecules stop at the TIa site. Perhaps, following cleavage in the presence of TFIIS, fifty percent of those ternary complexes elongating the shortened transcripts stop once again at the TIa site even in the presence of TFIIS and nucleotides, while the other fifty percent pass through to the run-off. 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 (Reines et al., 1987; T. Kerppola and C. Kane, unpublished results). On the other hand, in the presence of TFIIS cleavage is very rapid (Reines, 1992; Reines et al., 1992), and elongation through these sites can be nearly quantitative (Reines et al., 1992; Reines et al., 1989; SivaRaman et al., 1990). 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 (Gu et al., 1993; Krummel and Chamberlin, 1992a, 1992b; Lee and Landick, 1992; Linn and Luse, 1991; Rice et al., 1993). 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 (Kerppola and Kane, 1991; Spencer and Groudine, 1990). 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* (C. Kane, unpublished results). More than one mechanism promotes terminator read-through by elongation complexes of bacterial RNA polymerase (Turnbough and Landick, 1992), 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 (Surratt et al., 1991), and this cleavage reaction is promoted by at least two protein factors, GreA and GreB (Borukhov et al., 1992, 1993). Stalled ternary complexes formed with mammalian (Reines, 1992; Reines et al., 1992; Izban and Luse, 1992; Wang and Hawley, 1993), *Drosophila* (Guo and Price, 1993), and yeast

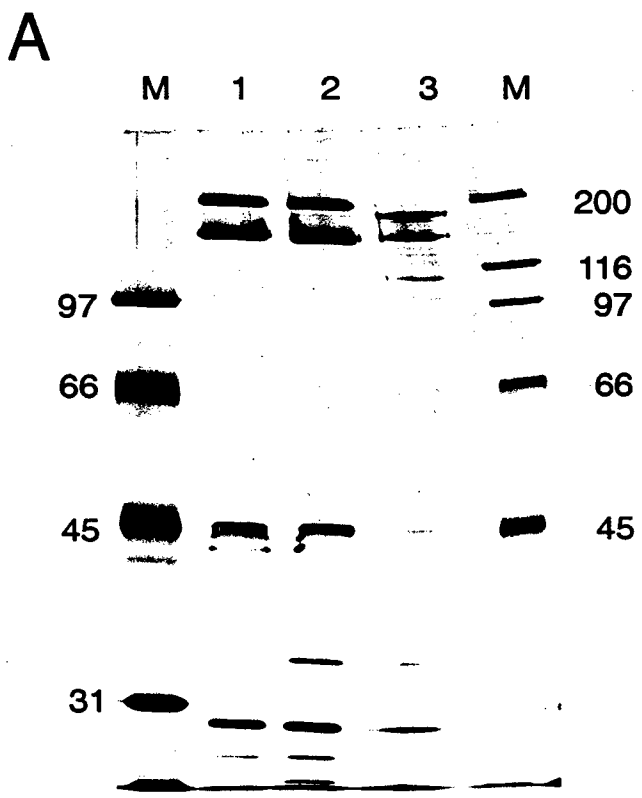
RNA polymerase II (this report) can cleave and then elongate nascent transcripts in response to TFIIS. Cleavage of the nascent transcript has also been observed in stalled ternary complexes formed with vaccinia RNA polymerase which contains a subunit with sequence similarity to TFIIS (Hagler and Shuman, 1993; Ahn et al., 1990). Cleavage of the nascent transcript has also been seen in isolated ternary complexes containing yeast RNA polymerase III (Whitehall et al., 1994). Further, 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 ten nucleotides (see Fig. 3 in Kuhn et al., 1990); 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 (Kuhn and Grummt, 1989), 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.

FIGURE LEGENDS

Figure 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 yeast RNA polymerase II, form Δ4,7. Lane 2: 1 μg yeast RNA polymerase II, form IIa. Lane 3: 0.5 μg 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 BioRad, sizes in kDa 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 truncated (D1-113) TFIIS. Lane 2: 0.25 μg full length TFIIS. Lane 3: 1.2 μg protein from partially purified TFIIS/P37-containing fraction, used in the experiment of Figure 2. Marker lane contains Silver Stain Low Molecular Weight Protein Standards from BioRad, sizes in kDa as indicated.


A
M
1
2
3
M
200
116
97
66
45
31
97
66
45

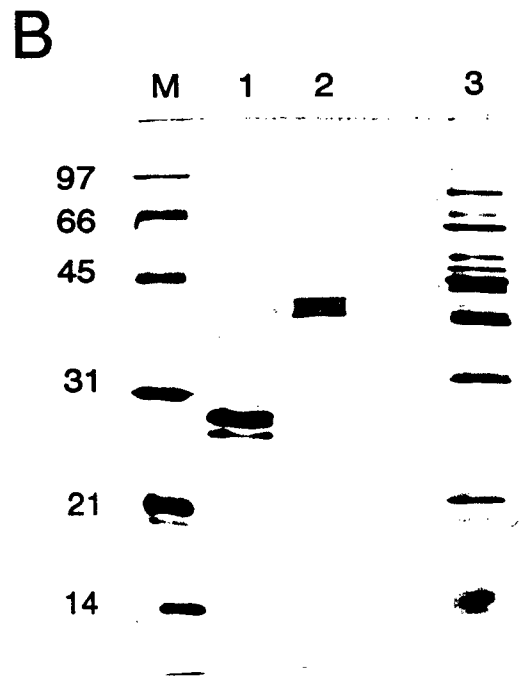

B
M
1
2
3
97
66
45
31
21
14

Figure 2 - Sequence of the Tailed Template pCpGEMTerm

The plasmid pGEMTerm was generated by insertion of the TaqI fragment of the human histone H3.3 gene from the first intron into the AccI site of the vector pGEM2. To generate the template pCpGEMTerm, the unique SmaI site was cut and terminal transferase was used to add 3'-deoxycytidine tails. PstI was used to generate a run-off transcript of a convenient size. EcoRI was used to prevent transcription from the other tail. Only the nontemplate strand is shown; the deoxycytidine extension is not shown since it is on the template strand. The SI mapped termination sites at the TII, TIb, and IIa sites are underlined (Reines et al., 1987). The box indicates the minimal sequence of the TIa site required to cause RNA polymerase II to pause (Kerppola and Kane, 1990).

Sequence of pCpGEMTerm

GGGGATCCTC TAGAGTCGAG CGGGAAAGGG GTGGAAATCG CCGCCGTCGC ACCCTGGGGT AACTCGCTTT
3maI TaqI
70

TII
TTGCTGCCTC CCCACAGTTG TTGGTAAGCT TCACCATCTT GTCTCTCTTC TCTGGTCACC GACCATATTT
140

TII TIb TIa
TTTCCCCCGT TTCTCTCTTC CCCATTTCAA AAAAGCAAGA ATTTTTAAAA GAGGGACGTT TTTTTCCCTT
210

TTTTGGAGAG GCGGAAACTT GATGCATTTG AAATGCAAAA AAAACTTTTG CATTTTGGAG GGCGGCACTG
280

RO
GCAGGGGATA GGGGTATTTT CGACCT
TaqI PstI
306

Figure 3 - 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 in Methods. An aliquot was removed and incubated with TFIIS storage buffer. The partially purified TFIIS/P37 containing fraction (300 ng 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.

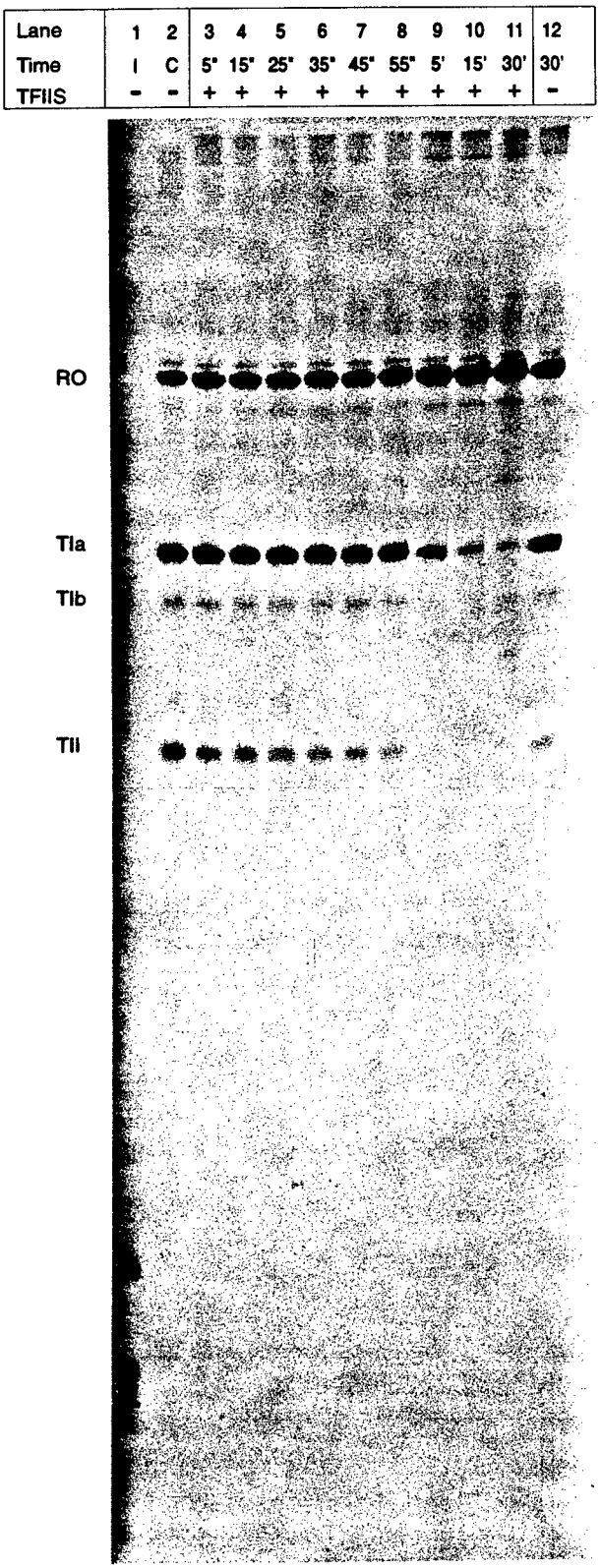
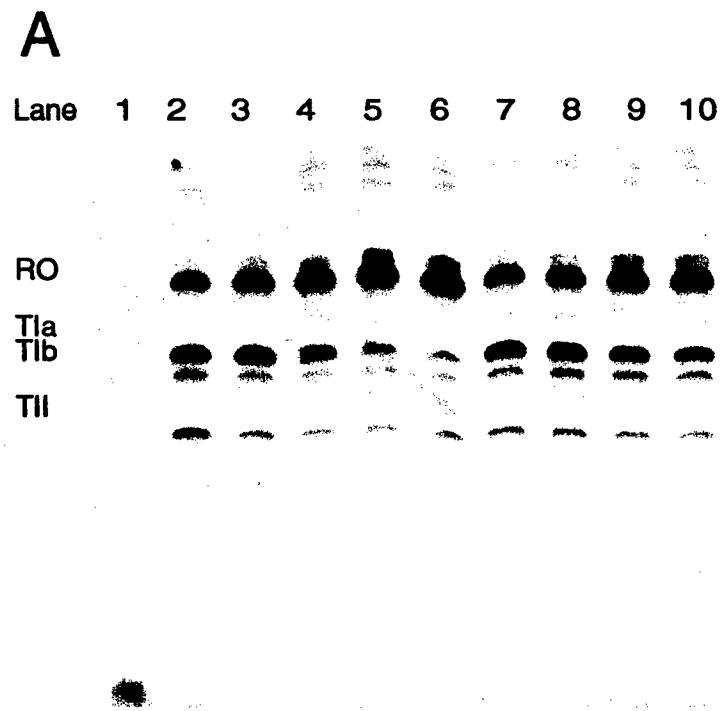

Lane 1 2 3 4 5 6 7 8 9 10 11 12
Time I C 5" 15" 25" 35" 45" 55" 5' 15' 30' 30'
TFIIS - - + + + + + + + + + -
RO
TIa
TIb
TII

Figure 4 - 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 in Methods. At 2 minutes 30 sec, 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 run off; Lanes 3-6: RNA from ternary complexes incubated with TFIIS, stopped after 5, 10, 30, or 50 minutes of incubation, respectively; Lanes 7-10: RNA from ternary complexes incubated in the absence of TFIIS, stopped after 5, 10, 30, or 50 minutes of incubation, respectively.


A
Lane 1 2 3 4 5 6 7 8 9 10
RO
TIa
TIb
TII

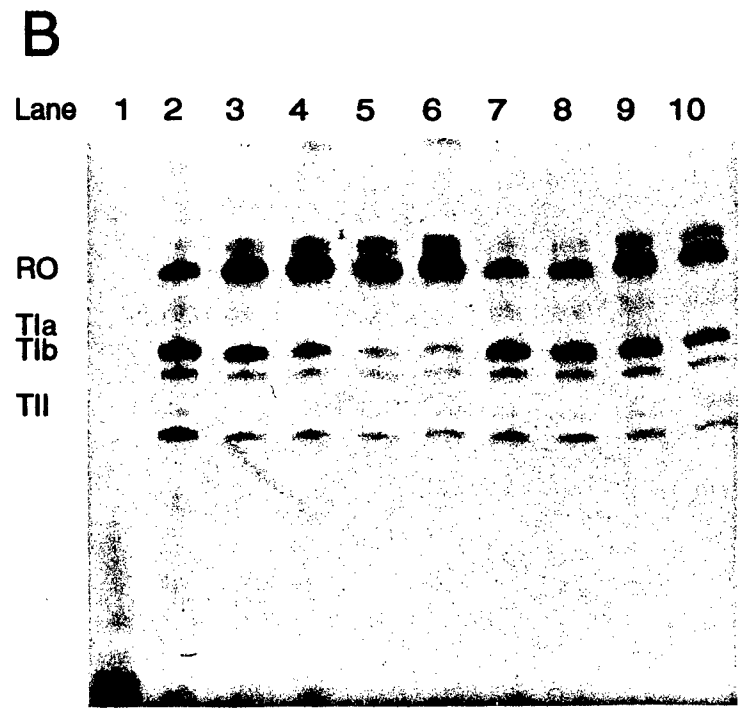

B
Lane 1 2 3 4 5 6 7 8 9 10
RO
TIa
TIb
TII

Figure 5 - 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 in Methods. Either full length (A) or truncated (B) TFIIS was added and reactions were incubated for the times indicated. RNAs were quantitated as described in Methods. **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 timepoints in the absence of TFIIS; open symbols represent timepoints in the presence of TFIIS.

△ = IIa. □ = Δ4,7. ○ = ΔCTD.

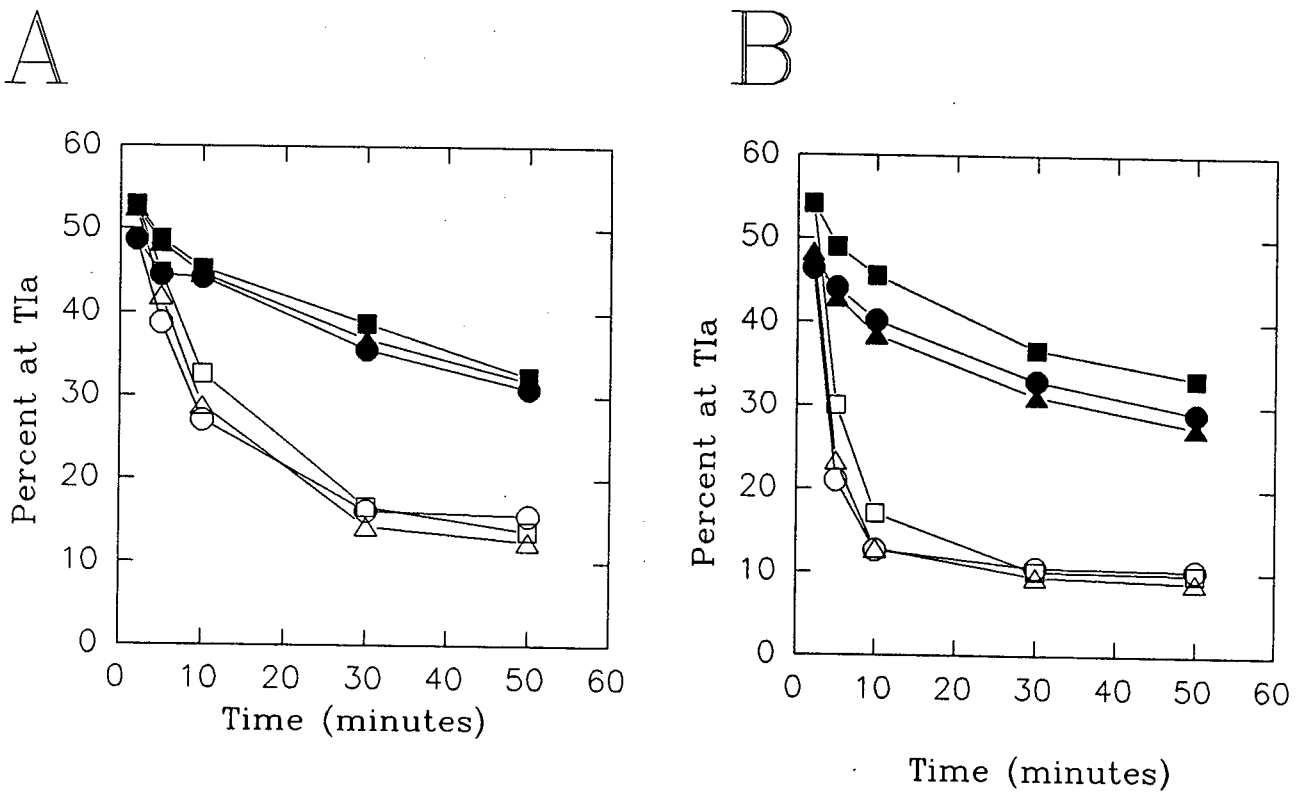
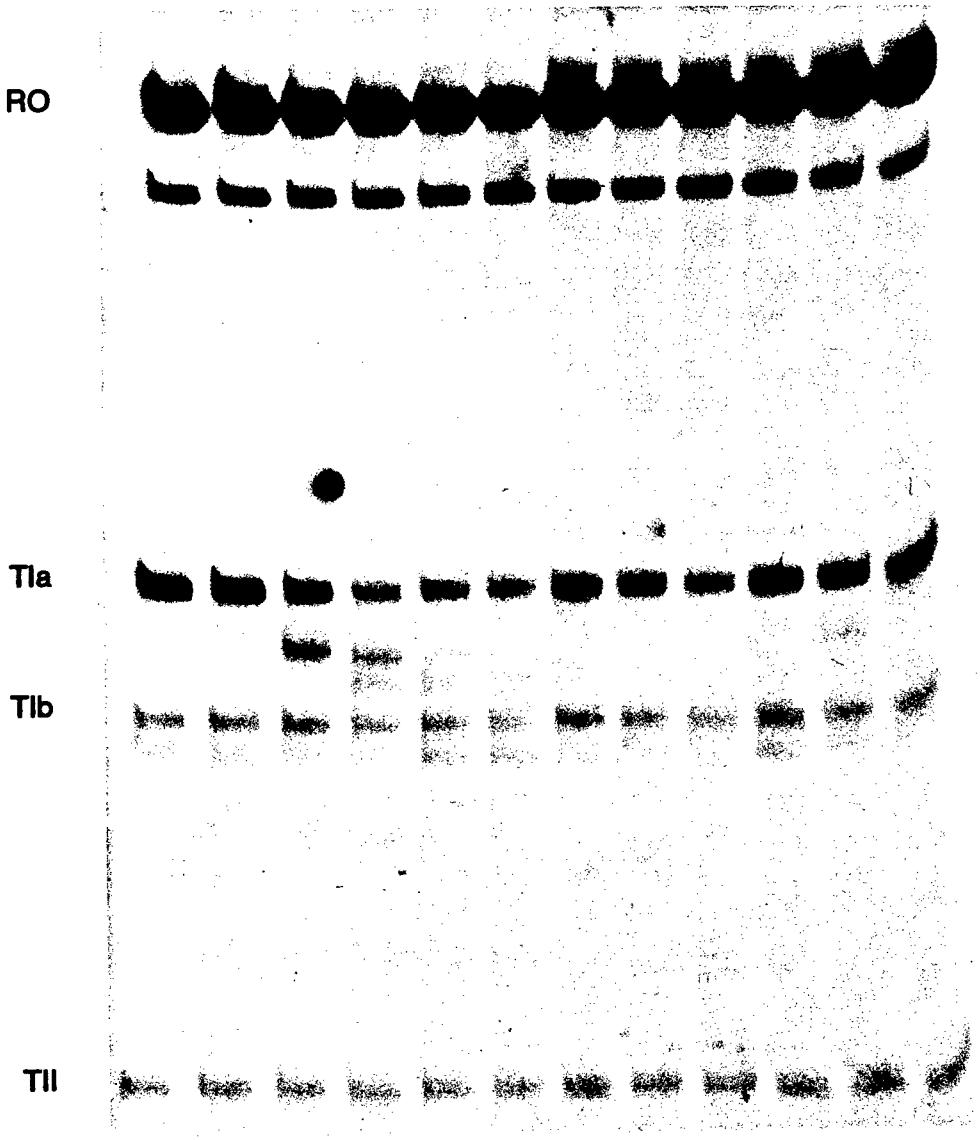

A
B
Percent at TIa
Time (minutes)
0
10
20
30
40
50
60

Figure 6 - TFIIS-Stimulated Cleavage of Nascent RNA in Ternary Complexes

Isolated ternary complexes were formed with RNA polymerase IIa as described in Methods. 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 seconds, 1 minute, 5 minutes, 10 minutes, or 20 minutes after TFIIS addition; Lanes 7-9: Nucleotides (0.8 mM) were added to complexes that had been incubated with TFIIS for 5 minutes 30 sec. Aliquots of the reaction were stopped at 10 seconds, 1 minute, or 20 minutes after nucleotide addition. Lane 10: RNAs from isolated ternary complexes (Lane 1) incubated 20 minutes with no further addition. Lanes 11 and 12: α-amanitin was pre-incubated with isolated ternary complexes for 2 minutes prior to the addition of TFIIS. An aliquot of the reaction was stopped 20 minutes after TFIIS addition (Lane 11). Nucleotides were added to the remainder and stopped after two minutes further incubation (Lane 12).

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	-	-	-	-	-	-	-	-	-	-	+	+



Chapter 3

Further Deletion Analysis of the Amino-terminus of Yeast TFIIS, *in vitro* and *in vivo*

INTRODUCTION

The regulation of transcriptional elongation is one of the myriad of mechanisms available for cells to control the expression of genes. In the last few years the number of genes thought to be regulated, at least in part, by control of transcriptional elongation has increased significantly (see Table 1 in Chapter 1). The still small size of this list probably reflects the difficulty of identifying blocks to transcriptional elongation *in vivo* rather than the rarity of this regulatory mechanism.

Although it is clear that many genes are regulated at this step of the transcription cycle, several different mechanisms seem to be involved (Krumm et al., 1993, 1995; Kane, 1994; Maldonado and Reinberg, 1995; Zawel and Reinberg, 1995; Kibel et al., 1995; Krumm and Groudine, 1995; Duan et al., 1995; Shor et al., 1995; Grossman and Tindall, 1995; Lee et al., 1992; Rasmussen and Lis, 1995; Jones and Peterlin, 1994). In some cases DNA sequences may be involved; in others protein factors seem to be required. Some blocks may be set up both by the DNA sequence and by DNA-binding proteins. However, for most genes in which blocks to elongation occur, the nature of the block is not understood; nor is the mechanism which regulates the ability of polymerase to read-through such sites.

Recently it has become apparent that the initiation reaction can play a significant role in determining the elongation properties of RNA polymerase. Some promoters, such as c-myc and hsp70, can direct the establishment of transcription complexes that are incompetent to pass through blocks to elongation, while others dictate elongation competent polymerase complexes (Rougvie and Lis, 1990; Krumm et al., 1993, 1995; Maldonado and Reinberg, 1995).

Several transcription elongation factors for RNA polymerase II have been identified and characterized *in vitro*, including TFIIS (Sekimizu et al., 1976; Sawadogo et al., 1979; Rappaport et al., 1987; reviewed in Kassavetis and Geiduschek, 1994; Reines, 1994; Kane, 1994), TFIIF (Finkelstein et al., 1992; Chang et al., 1993; Tan et al., 1995; Wang and Burton, 1995; reviewed in Conaway and Conaway, 1993), Elongin (Kibel et al., 1995; Duan et al., 1995; Aso et al., 1995; Bradsher et al., 1993a, 1993b), and p-TEFa and P-TEFb (Marshall and Price, 1992, 1995). Although TFIIS is probably the best characterized of these factors, very little is known about its function in cells. The general initiation factors TFIIE and TFIIH can also affect the elongation properties of polymerase (Maldonado and Reinberg, 1995; Goodrich and Tjian, 1994) and may possibly play an important role in determining whether elongation incompetent or competent complexes are initiated from promoters such as c-myc (M. Groudine, personal communication). Although these two general initiation factors are proposed to control promoter clearance, there is currently no information to connect the promoter specific effects on elongation with any of the other elongation factors. Yet it is clear from the many genes shown to contain regulated blocks to transcriptional elongation that the actions of factors within the cell contribute to this regulation. For example, the von Hippel-Lindau tumor suppressor protein was recently identified as a regulator of Elongin (Kibel et al., 1995; Duan et al., 1995; Aso et al., 1995; Krumm and Groudine, 1995). The mechanisms by which elongation factors may be targeted to the correct genes *in vivo* remain unknown. Better understanding of the transcription initiation reaction and its effect on the elongation reaction may provide clues.

Elucidation of the interactions of elongation factors with RNA polymerase itself will also provide information on how the elongation reaction may be regulated. In the previous chapter, I described characterization of the interaction of yeast TFIIS with RNA polymerase II. *In vitro*, TFIIS proteins stimulate RNA polymerase to cleave the nascent RNA transcript and to read-through intrinsic blocks to elongation. I demonstrated that an amino-terminal truncation of the TFIIS protein lacking the first 113 amino acids, referred to as Δ2-113 is indistinguishable from the full length protein when either cleavage or read-through activities are assayed.

Here I will describe an extension of the truncation analysis. Two additional truncation mutants of yeast TFIIS, Δ2-130 and Δ2-143 were analyzed *in vitro* for the ability to stimulate stalled ternary elongation complexes containing yeast RNA polymerase II to cleave the nascent transcript and to read-through an intrinsic block to elongation. Both were fully active compared to the wild type protein.

These three truncation mutants of yeast TFIIS were also assayed for activity *in vivo*. Since deletion of the *PPR2* gene encoding TFIIS confers sensitivity to the drug 6-azauracil, each of the truncation mutants was tested for the ability to complement the *in vivo* phenotype of the deletion. Any of the three truncated TFIIS proteins was sufficient to restore resistance to the drug. One additional truncation protein, Δ2-151, did not complement, a result consistent with a previous report that overproduction of a yeast Δ3-147 TFIIS protein can complement but a Δ2-149 cannot (Nakanishi et al., 1995). These analyses define the minimal region of the TFIIS protein required for productive interaction with RNA polymerase II and provide a basis for further study into the protein:protein interactions between the TFIIS protein and the polymerase.

In the process of constructing the necessary plasmids for *in vivo* expression, complete sequences of the regions upstream and downstream of the *PPR2* gene were obtained. Computer analysis of the promoter region identified putative binding sites for regulatory factors *ABF1, GRF2,* and *GCN4*. These will be discussed briefly although their relevance remains unknown.

MATERIALS AND METHODS

Materials

Restriction enzymes were obtained from New England Biolabs (NEB), Pharmacia, Boehringher Mannheim Biochemicals (BMB), or Bethesda Research Laboratories (BRL), as indicated. T4 DNA Ligase was obtained from United States Biochemicals (USB). Acetylated bovine serum albumin (BSA) was obtained from NEB. AmpliTaq DNA polymerase was obtained from Perkin Elmer-Cetus. Vent DNA polymerase was obtained from NEB. Calf Intestinal Phosphatase was obtained from NEB. 3'-deoxynucleotides were obtained from Pharmacia. Glycerol was obtained from Fluka. Gene Clean was obtained from Bio101. Polyester aquarium filter floss was obtained from Pet Food Express. 6-azauracil was obtained from Aldrich. Yeast nitrogen base without amino acids was obtained from Difco. Agar was obtained from Bacto. Primers for polymerase chain reaction (PCR) or sequencing were obtained from the Barker Hall DNA Synthesis Facility. Donkey anti-rabbit IgG antibody coupled to horse radish peroxidase was obtained from Amersham. ECL reagents were obtained from Amersham. Inhibit-ACE was obtained from 5'-->3'. Nucleotides were obtained from Pharmacia. [α-^{32}P]-CTP, >3000 μCi/mmole was obtained from Amersham or NEN. Bio-Gel 30 columns were obtained from BioRad.

RNA polymerase IIa was the generous gift of Aled Edwards (McMaster University, Ontario, Canada) and was purified by immunoaffinity as described (Thompson et al, 1990; and Edwards et al, 1990). Purified histidine-tagged preparations of variant yeast TFIIS proteins [His-yTFIIS(1-309), His-yTFIIS(Δ2-113), His-yTFIIS(Δ2-130), and His-yTFIIS(Δ2-143)] and the plasmids pET11aPPR2(1-309), pET11aPPR2(114-309), pET11aPPR2(131-309), and pET15bPPR2(144-309) were the

generous gift of Donald E. Awrey and Aled M. Edwards (McMaster University, Ontario, Canada).

Bacterial and Yeast Strains

Bacterial and yeast strains used and relevant parent strains are listed in Table 1.

Ligation Conditions

Ligation reactions were performed in 40 mM Tris-Cl, pH 7.5; 10 mM $MgCl_2$; 10 mM dithiothreiotol; 50 µg/ml acetylated BSA (NEB); 2 mM ATP and 1 unit T4 DNA ligase (USB) in a volume of 20 µl. Generally, as much insert DNA as possible was used, with an amount of vector that could be visualized by running a quarter of the ligation reaction on a 1% agarose gel. When possible, ligations were performed at multiple insert to vector ratios by varying the amount of insert DNA added to a constant amount of vector DNA.

Transformation of *Escherichia coli*

The strain of *E. coli* used for all transformations of ligation reactions (except that of pKC6) was DH10B (Grant et al, 1990). These cells were made competent by a calcium-manganese protocol of D. Hanahan (Hanahan et al, 1991). Briefly, cells were streaked on to a fresh LB plate from a stab and incubated at 37°C overnight. Several colonies were picked into 1 ml SOB - Mg (2% Bacto tryptone, 0.5% Bacto yeast extract, 10 mM NaCl, and 2.5 mM KCl), vortexed to disperse the cells, and used to inoculate 100 mls SOB without magnesium, in a 1L Erlenmeyer flask. This culture was shaken at 30°C until the OD_{550} reached 0.3, (generally 4 to 5 hours). The cells were transferred to a sterile GS-3 bottle, chilled on ice for 10

minutes, and pelleted by spinning 10 minutes at 2500 rpm in a Sorvall GS-3 rotor. The supernatant was discarded and the cells were resuspended gently in 1/3 volume (35 mls) cold CCMB-80 [10 mM KOAc; 10% glycerol (Fluka); 80 mM $CaCl_2$; 20 mM $MnCl_2$; 10 mM $MgCl_2$; pH 6.4]. The cells were incubated on ice 20 minutes and then pelleted as above. The cell pellet was resuspended in 8 mls cold CCMB-80. At this point cells are ready for transformation immediately or could be aliquotted and frozen in a dry ice/ethanol bath. In either case, transformation efficiency was generally around 5 x 10^6 colonies/μg DNA and did not decrease much with storage at -80°C. This transformation efficiency was lower than reported by the published protocol (by 1-2 orders of magnitude), but was sufficient for transformation of ligation reactions.

For transformation, cells were allowed to thaw on ice, aliquotted 200 μl per Eppendorf tube, and incubated on ice 10-20 min. DNA was added in no more than 20 μl, mixed, and the cells were incubated on ice for 20 minutes more. Heat shock was for 90 sec at 42°C and cells were then placed on ice for approximately 2 minutes to cool. 800 μl of LB or SOC (SOB - Mg + 20 mM $MgCl_2$, 20 mM $MgSO_4$, and 0.4% glucose) was added and the cells were shaken 1 hour at 37°C. 100 μl of each transformation mix was plated on an LB-amp plate, the remainder on a 2nd plate, and the plates were incubated overnight at 37°C.

Lithium Chloride / Triton X-100 DNA Preparations from *E. coli*

DNA from *E. coli* was prepared using a lithium chloride/Triton X-100 mini-preparation (Hofmann and Brian, 1991; Goode and Feinstein, 1992). 2 ml LB-ampicillin cultures were inoculated with *E. coli* containing a plasmid and shaken overnight at 37°C. Cells were harvested by centrifugation in Eppendorf tubes and

the supernatant was discarded. To each tube was added 100 µl LiCl/Triton solution [2.5 M LiCl; 50 mM Tris-Cl, pH 8.0; 4% Triton X-100; 62.5 mM EDTA], 50 µl neutralized phenol, and 50 µl SEVAG. These mixtures were vortexed vigorously for 10 seconds and spun 2 minutes in a microfuge. The cell debris formed a compact pellicle above the organic layer. The supernatant was transferred to a new tube containing 200 µl cold ethanol and the DNA was precipitated by spinning at least 15 minutes at 4°C in a microfuge. The pellet was washed with 70% ethanol, dried and dissolved in 20 µl H_2O.

Alkaline Lysis Preparations of Plasmids Suitable for Sequencing

Generally, *E. coli* harboring the desired plasmids were inoculated into 100 mls of LB containing 50 µg/ml ampicillin. The cultures were shaken vigorously at 37°C overnight. Cells were harvested by spinning 5 minutes at 5K rpm in either a GSA or a GS-3 Sorvall rotor. The cell pellet was resuspended in 8 mls GTE [1% glucose, 25 mM Tris-Cl, pH 8, 10 mM EDTA]. After the addition of 16 mls 0.2 M NaOH/1% SDS, the mixture was swirled and then incubated on ice for 5 min. Next, 8 mls of KOAc [3 M in K^+, 5 M in OAc^-] was added, the mixture was rolled gently to mix to avoid shearing chromosomal DNA, and incubated on ice 15 min. Cell debris was pelleted by spinning 10 minutes at 6K rpm in a Sorvall GS-3 rotor. The supernatant was filtered through cheesecloth into a clean GSA bottle and 0.6 volumes of isopropanol was added. The DNA was precipitated by spinning 7 minutes at 6K rpm in a Sorvall GSA rotor. The pellets were allowed to air dry while inverted for 10-20 minutes and then dissolved in 0.4 ml TE. Usually RNaseA was added to about 30 µg/ml and incubated 5 minutes at 37°C. RNaseA and other proteins were removed by phenol/SEVAG extraction. The DNA was precipitated

with 0.3 M NaOAc and ethanol and the pellet was dissolved in 500 µl H_2O. When prepared from DH10B *E. coli* cells, this DNA was used without further preparation required for subcloning or sequencing.

Purification of Oligomers by Polyacrylamide Gel Electrophoresis

Long DNA oligomers were gel purified using a 15% acrylamide/8.3M urea preparative gel, run at 35W for 5-8 hours (until the zylene cyanol migrated 14-20 cm). The bands were visualized using short wave ultraviolet (UV) with fluorochrome-saturated thin-layer chromatography (TLC) plates behind the Saran-wrapped gel. The longest product bands of each oligomer were excised, homogenized with a glass pestle, and extracted with 0.2x TE overnight on a rotating wheel. The acrylamide was pelleted by spinning for 10 minutes at speed 5 in an IEC clinical centrifuge. The supernatant was transferred to a new tube and the acrylamide pellet was washed again as above. The supernatants were combined and extracted with butanol until the volume was reduced to approximately 0.5 ml. The remaining volume of the oligomers were extracted one time with phenol/chloroform and one time with chloroform. The oligomers were precipitated with 0.3M NaOAc, 10 mM $MgCl_2$, and ethanol. The pellets were dried and dissolved in 200 µl H_2O.

Optimization of PCR with Vent DNA Polymerase

For each different combination of template and primers, PCR was optimized for Vent DNA polymerase (NEB) following the guidelines provided by the manufacturer. The extension time was determined according to the formula of 1 minute per 1 kilobase (kb) in length of the product. The optimal annealing

temperature and magnesium sulfate concentration were determined empirically. A set of three 100 µl reactions were set up using standard conditions for Vent [20 mM Tris-Cl, pH 8.8; 10 mM KCl; 10 mM $(NH_4)_2SO4$; 2 mM $MgSO_4$; 0.1% Triton X-100; 0.4 µM dNTPs; 10 µM each primer; and a nanogram or less template]. Additional $MgSO_4$ was added to two of the reactions to bring them to final concentrations of 4 mM $MgSO_4$ or 6 mM $MgSO_4$. To each reaction, 2 units of Vent enzyme was added. The first such set was amplified using this cycle: Denaturation at 94°C for 1 minute; Annealing at 47°C for 1 minute; and Extension at 72°C for the time determined by the formula above, for 25 cycles. For subsequent sets, the annealing temperature was raised or lowered by 2°C. This process continued until an optimal reaction condition was determined.

Preparative PCR Using Vent DNA Polymerase

Preparative PCR using Vent DNA polymerase was done on scale of 1 ml. A single 1 ml mixture containing all reagents except the enzyme was prepared with the optimal magnesium concentration; 240 µl was aliquotted to each of 4 thin-walled PCR tubes. 4 units Vent were added to each tube. Amplification was performed according to the optimal cycling as determined above. The 4 tubes were pooled and extracted with phenol/SEVAG. The DNA was precipitated with 0.3 M NaOAc and ethanol. The pellet was dried and resuspended in 100 µl H_2O.

Construction of pKC6

pKC3 DNA was digested with the blunt-cutting enzymes EcoRV (NEB) and SmaI (NEB), which flank the EcoRI site within the pBS.KS+ polylinker in pKC3 (see Chapter 2 Methods for construction). DNA was phenol extracted, ethanol

precipitated, and dissolved in 10 µl H_2O. The DNA was treated with T4 DNA ligase and incubated at 25°C overnight. The ligation mixture was electroporated into XL-1Blue *E. coli* cells. DNA was prepared from transformants and restriction analysis indicated that in one transformant the polylinker EcoRI site had been removed. Fortuitously, the BamHI site proximal to the EcoRV, EcoRI, and SmaI sites had also been destroyed. This plasmid was designated pKC6. The insert contained in pKC3 and in pKC6 is diagrammed in Figure 2; sites destroyed in the construction of pKC6 are indicated.

Construction of pKC-RS314

A new polylinker, referred to as KRC-POLY, was constructed using the polymerase chain reaction (PCR) and two long, partially self-complementary oligomers, KRCPOLY1 and KRCPOLY2. The sequence of the KRCPOLY1 oligomer (79 bases) is 5'-CGGGGTACCGGGCCCATCGATGAATTCCTGCAGGTCTACCCG-GGCGGCCGCACTAGTCTCGAGCATTCATCCATATGGC-3'. The sequence of the KRCPOLY2 oligomer (73 bases) is 5'-TCCGGAGCTCCCGCGGAAGCTTATGCA-TAGATCTAGAGGATCCCGATCGGCATGCCATATGGATGAATGCTCG-3'. Both oligomers were gel purified as described above. The conditions for PCR with AmpliTaq DNA polymerase were 20 mM Tris-Cl, pH 8.5; 5 mM MgCl2; 50 mM KCl; 0.1% Triton-X100; 0.8 mM deoxynucleotides (dNTPs); 1 µM each primer; and 2.5 units AmpliTaq (Perkin-Elmer Cetus) per 100 µl reaction, with a cycle of denaturation at 94°C for 1 minute; annealing at 47°C for 1 minute; and extension at 72°C for 2 minutes for 25 cycles.

The PCR product KRC-POLY and the plasmid pRS314 (Sikorski and Hieter, 1989) were digested with the enzymes ApaI (NEB) and SacII (NEB), which cut at

either end of both KRC-POLY and the pBS.KS+ polylinker of pRS314. Doubly cut pRS314 DNA was gel purified using a 1% agarose gel and extracted from the agarose using Gene Clean (Bio101). Doubly cut KRC-POLY DNA was gel purified using a 2% agarose gel and extracted from the agarose by centrifugation through polyester filter floss (Glenn and Glenn, 1994). The two ApaI/SacII digested DNAs were ligated using T4 DNA ligase at 16°C. Ligation products were transformed into DH10B *E. coli* cells using the calcium-manganese protocol described above. DNA from transformants was prepared and restriction analysis identified a plasmid with the desired insert. Further restriction analysis and sequence analysis confirmed that the plasmid contained all of the necessary restriction enzyme sites; it was designated pKC-RS314.

Construction of pKC15

The 3'-flanking region of the PPR2 gene was amplified from the plasmid pKC6 using PCR with Vent DNA polymerase and the primers P37-15D3' and the polylinker primer T7+ (Figure 2). The sequence of T7+ is 5'-GCGTAATACGA-CTCACTATAGG-3'. The sequence of P37-15D3' is 5'-CGGGATCCAAAAGCCTGTT-CTGACTCTATTCG-3'. This primer incorporates a BamHI site for cloning and anneals slightly downstream of the end of the PPR2 open reading frame (ORF). The optimal conditions for this primer-template combination using Vent included 2 mM $MgSO_4$, a 45°C annealing temperature, and an extension time of 1 minute 30 sec. DNA for subcloning was generated with a preparative reaction at the optimal conditions as described above.

pKC-RS314 DNA was digested with BamHI (NEB) until digestion appeared complete as determined by the appearance of a 1 µl sample of the 50 µl reaction

on a 1% agarose gel. This DNA was then treated with Calf Intestinal Phosphatase (CIP). The reaction was diluted to 200 µl with 1X Universal Buffer (Stratagene), 1 unit of CIP (NEB) was added, and the reaction was incubated at 37°C for 1 hour. The reaction was stopped by the addition of EDTA to 2.5 mM and the phosphatase was removed by phenol/SEVAG extraction. The DNA was precipitated and resuspended in 30 µl H_2O.

The 3'-flanking sequence generated by PCR was digested with BamHI (NEB), which cuts the site added with the primer P37-15D3' and in the pKC6 polylinker, and with BglII (Pharmacia), which cuts a yeast genomic site approximately 700 base pairs downstream of the *PPR2* ORF. The enzymes were removed by phenol/SEVAG extraction. The DNA was precipitated and resuspended in 30 µl H_2O.

These DNAs, pKC-RS314/BamHI/CIP and 3'-PCR-product/BamHI/BglII, were ligated using T4 DNA ligase at 16°C and at 25°C. A sample of each ligation was resolved on a 1% agarose gel to visualize the products and revealed significant increases in mobility relative to the starting DNAs, regardless of the incubation temperature. Ligation products were transformed into DH10B *E. coli* cells using the calcium-manganese protocol described above. DNA was prepared from transformants and restriction analysis indicated that one plasmid contained an insert of the expected size. Further restriction analysis and PCR indicated that the insert was in the correct orientation. Sequence analysis of this plasmid indicated that there were no mutations in the PCR-generated 3'-flanking sequence. In addition, the sequence indicated that a portion of KRC-POLY on the downstream side (with respect to the direction of transcription of the PPR2 gene) of the insert had been lost and replaced with the pBS.KS+ polylinker fragment, presumably

generated by BamHI cleavage within the pKC6 polylinker sequence of the PCR product. As these unexpected alterations of the polylinker were inconsequential for subsequent subcloning steps, this plasmid was designated pKC15 (Figure 3).

Construction of pKC16

The 5'-flanking region of the PPR2 gene was amplified from the plasmid pKC6 using PCR with Vent DNA polymerase and the primers T3+ and P37-14U5' (Figure 2). The sequence of T3+ is 5'-GCAATTAACCCTCACTAAAGGG-3'. The sequence of P37-14U5' is 5'-AACTGCAGCTCGAGTGCGGACTGACTACACTG-3'. The primer is located just upstream of the beginning of the PPR2 ORF and incorporates an XhoI for cloning. The optimal condition for this primer-template combination using Vent used 2 mM $MgSO_4$, a 53°C annealing temperature, and an extension time of 1 min. DNA for subcloning was generated with a preparative reaction at the optimal conditions as described above.

pKC15 DNA and PCR-generated DNA encoding the 5'-flanking sequence of the PPR2 gene were digested with the enzymes ClaI (BRL) and XhoI (NEB). The enzymes were inactivated by incubating 20' at 65°C. These DNAs were ligated using T4 DNA ligase at 16°C at three different insert to vector ratios. A sample of each ligation was resolved on a 1% agarose gel to visualize the products and revealed significant increases in mobility relative to the starting DNAs. Ligation products were transformed into DH10B *E. coli* cells using the calcium-manganese protocol described above. DNA was prepared from 40 transformants and plasmids containing an insert were distinguished from pKC15 by a shift in supercoiled mobility. Restriction analysis of two of these transformants confirmed the presence of the appropriate insert. Sequence analysis of one of these plasmids indicated that

there were no PCR-generated mutations in the promoter region between the start codons of the divergently-transcribed ORF's, *PPR2* and *RNA15*, as discussed below. The RNA15 ORF was not completely sequenced in this plasmid, designated pKC16 (Figure 3).

Construction of pKC16 derivatives containing PPR2 ORF's

DNAs for subcloning the various PPR2 ORF's were generated by PCR. The template for the full length (309 amino acids) ORF was pET11aPPR2(1-309). The templates for three of the truncated ORF's were pET11aPPR2(114-309), pET11aPPR2(131-309), and pET15bPPR2(144-309) where the numbers in parentheses indicate the amino acids of the PPR2 ORF (numbered according to the full length protein) encoded by the insert following the ATG added by the NdeI site. These four variant PPR2 ORF's were amplified from the plasmids listed above using PCR with Vent DNA polymerase and the vector primers T7+ and T7TERM. The sequence of T7+ was given above and the sequence of T7TERM is 5'-TGCTA-GTTATTGCTCAGCGGTGG-3'. The optimal conditions for these primer-template combinations using Vent utilized 2 mM $MgSO_4$, a 47°C annealing temperature, and an extension time of 1 min. DNA for subcloning was generated with a preparative reaction at the optimal conditions as described above.

One additional truncated PPR2 ORF was generated by PCR with pKC6 as the template, Vent DNA polymerase, and the primers P37-30D and P37-31U. The sequence of P37-30D is 5'-CGCGCACCATATGGTACTAAAAGCACTCTACGACG-3' and it incorporates an NdeI site for cloning such that it adds an ATG codon immediately before the codon for the 152nd amino acid of PPR2. The sequence of P37-31U is 5'-GCGAGCCGGATCCCCTTCCTCCTATTGTTTCCTT-3' and incorporates

a BamHI site for cloning. The primer anneals slightly downstream of the PPR2 ORF and is identical in sequence to the 3' primer described in Chapter 2. The optimal condition for this primer-template combination using Vent utilized 2 mM $MgSO_4$, a 51°C annealing temperature, and an extension time of 30°C. DNA for subcloning was generated with a preparative reaction at the optimal conditions as described above.

DNAs of the five PCR-generated fragments, each encoding a variant of the PPR2 ORF, and pKC16 were digested with the enzyme NdeI (NEB) until agarose gel analysis indicated complete cutting of each DNA. At this point, BamHI (NEB) was added to the digest and incubated until digestion appeared to be complete. The enzymes were removed by phenol/SEVAG extraction. The DNAs were precipitated with 0.3M NaOAc and ethanol and the pellets were dissolved in 30 µl H_2O, except for the pKC16 DNA, which was dissolved in 50 µl H_2O. These DNAs were ligated using T4 DNA ligase at 16°C at three different insert to vector ratios. A sample of each ligation was resolved on a 1% agarose gel to visualize the products and revealed significant increases in mobility relative to the starting DNAs. Ligation products were transformed into DH10B *E. coli* cells using the calcium-manganese protocol described above. DNA was prepared from 15 or 20 transformants of each different insert and plasmids containing an insert were distinguished from pKC16 by a shift in supercoiled mobility. Restriction analysis of transformants confirmed the presence of the appropriate insert. Sequence analysis of five plasmids, each with a different insert indicated that four of the five did not contain any PCR-generated mutations within the PPR2 ORF region. These were designated pKC16(1-309), pKC16(114-309), pKC16(144-309), and pKC16(152-309). The fifth plasmid contained one mutation, a C --> T change, resulting in a proline

to leucine mutation at position 292 of the protein. This plasmid was designated pKC16(131-309:P292L).

Sequencing and Sequence Analysis

The 3.9 kb insert of the plasmid pKC3 was sequenced in its entirety as a basis of comparison for DNAs cloned using PCR. See Table 2 for primers used. All fragments cloned by PCR techniques were sequenced entirely, with the exception of the insert of pKC16, as noted above. All sequencing was performed by the Barker Hall Sequencing Facility using the PRISM™ Ready Reaction DyeDeoxy™ Terminator Cycle Sequencing Kit with an ABI Model 373A DNA Sequencer. Sequences were aligned and edited using the Gel program of the IntelliGenetics Suite. Database searches to determine whether the 5' and 3'-flanking regions of the *PPR2* gene contained other known genes were performed using blastp and blastn (Altschul et al, 1990).

Transformation of Supercoiled Plasmids into *Saccharomyces cerevisiae*

S. cerevisiae cells were made competent for transformation by the protocol of Elble (1992). 10 ml cultures of YPD were inoculated with the desired strain and grown at 30°C, until mid-late logarithmic phase. Aliquots (0.5 ml) of cells were collected by centrifugation in Eppendorf tubes for 1 minute. The cell pellets were resuspended in the residual liquid and 10 µl of 10 mg/ml sheared salmon sperm DNA, the appropriate supercoiled plasmid DNA in a volume of 1 µl, and 0.5 mls of PEG/LiOAc/TE mixture (40 % PEG, MW 3350; 100 mM LiOAc, 10 mM Tris-Cl, pH 7.5; 1 mM EDTA) were added and mixed by vortexing. The cells were incubated at room temperature for 4-12 hours. Approximately 1 ml H_2O was added and

mixed into the cells/PEG mixture by inversion. The cells were pelleted by spinning 30 seconds in a microfuge. The supernatant was discarded, the cells were resuspended in 200 µl H_2O and spread on SC-tryptophan to select for cells receiving plasmids. Transformants were restreaked on SC-tryptophan, -uracil to select for the maintenance of the URA3 gene as well as for the plasmid.

Growth Media for Yeast Strains With or Without 6-azauracil

Solid media lacking or containing various concentrations of the drug 6-azauracil (6-AU) was prepared as follows. A 5 mg/ml solution of 6-AU (Aldrich) was prepared, adjusted to pH 7 with 1M NaOH, and sterilized by filtration. A solution of 40% dextrose was prepared and sterilized by autoclaving. SC-tryptophan, uracil was prepared in 0.5 L aliquots as follows: 3.35g Difco yeast nitrogen base w/o amino acids, 0.56g "almost complete" amino acid mix, 0.325g uracil drop-out mix, and 455 mls H_2O were mixed. ["Almost complete" amino acid mix and Uracil drop-out mix are described in Chapter 4 Methods.] For solid media, 10 g agar (Bacto) was added and the mixture was autoclaved and cooled to about 50°C. To the agar mixture was added, 25 mls 40% glucose and a total of 20 mls of H_2O and 5 mg/ml 6-AU to achieve the final 6-AU concentration desired. This mixture was poured into 100 mm Petri dishes (approximately 30 mls/dish) and allowed to cool. For liquid media, 25 mls 40% glucose and a combination of 20 mls of H_2O and/or 5 mg/ml 6-AU to achieve the final 6-AU concentration desired were added and the entire mixture was sterilized by filtration.

Test of Sensitivity of Yeast Strains to 6-azauracil on Solid Media

2 ml cultures of SC-tryptophan, uracil were inoculated with KJY1 containing one of the following plasmids: pRS314, pKC-RS314, pKC16, pKJ1, pKC16(1-309), pKC16(114-309), pKC16(131-309:P292L), pKC16(144-309), or pKC16(152-309), or with CMKY5 containing one of the plasmids listed above. Cultures were shaken at 30°C. A series of ten-fold dilutions in H_2O was prepared for each culture, from undiluted to 10,000-fold dilution. 400 µl aliquots of each dilution were placed into wells. Cells were replica plated onto plates with no drug and with the following concentrations of 6-AU: 10 µg/ml, 25 µg/ml, 50 µg/ml, 75 µg/ml, 100 µg/ml, 150 µg/ml, and 200 µg/ml and incubated at 30°C. All plates were photographed using T-MAX film at 3 and 4.5 days of growth.

Test of Sensitivity of Yeast Strains to 6-azauracil in Liquid Media

3 ml cultures of SC-tryptophan, uracil were inoculated with KJY1 containing one of the following plasmids: pKC16, pKJ1, pKC16(1-309), or pKC16(131-309:P292L) and shaken at 30°C. The OD_{600} of each culture was measured and the volume of each needed to inoculate a 40 ml culture to an OD_{600} of 0.01 was calculated. For each strain, a 40 ml SC-tryptophan, uracil culture and a 40 ml SC-tryptophan, uracil + 50 µg/ml 6-AU culture was inoculated with the volume calculated above. The initial OD_{600} was determined and the cultures were shaken at 30°C. Aliquots were removed at various times to determine the OD_{600} with a Beckman DU®-64 spectrophotometer (Figure 6).

Preparation of Protein Extracts from *S. cerevisiae*

Cultures of SC-tryptophan, uracil (20 mls) were inoculated with KJY1 containing one of the following plasmids: pRS314, pKC-RS314, pKC16, pKJ1, pKC16(1-309), pKC16(114-309), pKC16(131-309:P292L), pKC16(144-309), or pKC16(152-309) and shaken at 30°C. Aliquots were removed at various times to determine the OD_{600} of each culture. When cultures reached an OD_{600} of approximately 0.7 - 0.8, two OD_{600} units were pelleted in an Eppendorf. The supernatant was discarded, the cells were resuspended in 20 µl Sample Buffer (80 mM Tris-Cl, pH 6.8; 2% SDS; 0.1% bromophenol blue (BPB); 100 mM dithiothreiotol (DTT), 10% glycerol), and heated for 2 minutes at 95°C. Glass beads (0.2 g) were added to the sample, which was then vortexed for 1 minute at high speed. An additional 80 µl Sample Buffer was added to each and the samples were stored at -20°C until use. Protein extracts were vortexed briefly and heated 1 minute at 95°C prior to loading on SDS/polyacrylamide gels.

Western Blotting Analysis of Yeast Cell Extracts

Samples of proteins (15 µl each) from KJY1 cells bearing a plasmid expressing a TFIIS variant or one of the control plasmids, prepared as described above, were resolved on a 15% SDS/polyacrylamide gel. On the same gel, 100 ng samples of the four purified, histidine-tagged yeast TFIIS variants, His-yTFIIS(1-309), His-yTFIIS(Δ2-113), His-yTFIIS(Δ2-130), and His-yTFIIS(Δ2-143) were also resolved. Proteins were transferred from an SDS/polyacrylamide gel to a nitrocellulose filter overnight at 10V in Transfer Buffer (50 mM Tris; 380 mM glycine; 0.1% SDS; 20% methanol). The filter was stained in Ponceau S to visualize transfer, which appeared uniform. After transfer, the gel was stained in Coomassie

to visualize any proteins remaining in the gel. The staining indicated that although some of the higher molecular proteins had not transferred, transfer in the region of the full length and truncated TFIIS proteins was essentially complete. The filter was blocked by shaking in TBS-T (25 mM Tris-Cl, pH 7.5; 150 mM NaCl, 1% Tween-20) + 5% milk for 2 hours. Prior to binding of the primary antibody, the filter was washed 15 minutes in 50 mls TBS-T, and then 2 more times for 5 minutes each.

The primary antibody was a rabbit IgG raised against the full length yeast protein expressed in *E. coli* (C. M. Kane, unpublished results) and was used at 1:5000 dilution in TBS-T. A large piece of Parafilm (greater than twice the size of the filter) was creased in the middle. The filter was placed next to the right of the crease in the Parafilm. Two mls of the 1:5000 dilution of the antibody were pipeted onto the filter. The left side of the parafilm was gently laid over the filter, slowly to prevent trapping air bubbles, and the edges of the parafilm were sealed together around the filter, which was incubated with the primary antibody for 1 hr.

Before incubating with the secondary antibody, the filter was washed 2 times for 5 minutes in TBS-T + 5% milk and 1 time for 5 minutes in TBS-T; the series of three washes was repeated one time. The secondary antibody was donkey anti-rabbit IgG coupled to horse radish peroxidase (Amersham) also used at 1:5000 dilution. The filter was incubated with the secondary antibody for 1 hr in the manner described above. The filter was then washed 15 minutes in TBS-T and four more times for 5 minutes in TBS-T. Visualization was performed with Amersham's ECL kit.

SDS/Polyacrylamide Gel Electrophoresis

Preparations of Histidine-tagged yeast TFIIS (His-yTFIIS) proteins or protein extracts from *S. cerevisiae* cells were resolved on 15% polyacrylamide gels.

Silver Staining of SDS/Polyacrylamide Gels

This was done as described in Chapter 2

Read-Through Assays of His-yTFIIS Proteins

This was done as described in Chapter 2

Quantitation of Transcripts in Polyacrylamide Gels

Quantitation of radioactivity in polyacrylamide gels was performed using the Molecular Dynamics PhosphorImager system with ImageQuant software, version 3.3. Transcripts with 3' ends at TII, TIb, TIa and the run-off transcript were identified, and the cpm in each was determined using local background computation. The total cpm in these four types of transcripts was determined for each lane 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) - (TII cpm) - (TII cpm)].

Ternary Complex Cleavage Assay

This was done as described in Chapter 2.

RESULTS

Purpose

The major goal of this work was to determine if the amino terminus of the TFIIS protein is required for its activity *in vivo*, since it is not required for TFIIS activity *in vitro*. A series of amino-terminal truncation mutants of the yeast TFIIS protein were tested to determine how much of the amino-terminus was required for TFIIS function in cells. It has been known for some time that the C-terminal portion of TFIIS is sufficient for the known *in vitro* activities (Horikoshi et al., 1985; Agarwal et al., 1991; Ciprés-Palacín and Kane, 1994; Guo and Price, 1993; Christie et al., 1994; and see Chapter 2 of this report). As the amino-terminus of the mouse protein is phosphorylated *in vivo*, it has been proposed that the amino-terminus of TFIIS might be involved in the regulation of its activity in cells (Sekimizu et al., 1981). If the amino-terminus of the TFIIS protein is involved in regulating the activity of the protein, there are several possible mechanisms by which this could occur. This region of the protein might contain sites of post-translational modification; alternatively, it might interact with regulatory proteins. Either of these possibilities could modulate the interaction of TFIIS with RNA polymerase II or the degradation of the protein, to name only two possible regulatory strategies. Another possible function of the amino-terminus is nuclear localization. This region contains many basic residues including a match to a consensus nuclear localization signal (NLS; Boulikas, 1993, 1994). In this case, the amino-terminus would be dispensable *in vitro* but essential *in vivo*.

In addition to testing for function *in vivo*, the series of TFIIS truncation proteins was tested for activity *in vitro*, extending the work described in Chapter 2. The combination of *in vivo* and *in vitro* techniques provided a robust approach to

investigate the functional requirement for the amino-terminus of the TFIIS protein. The proteins tested included deletions of the first 113, 130, or 143 amino acids; the full length protein provided a control. The ability of each of these four proteins to stimulate transcript cleavage and read-through was determined. In order to test each of these proteins *in vivo*, a vector was constructed to express them from the natural promoter. The natural 3' untranslated sequences were also included to preclude altering any regulatory effects which might act on these sequences. Activity *in vivo* was assayed by determining the ability of the TFIIS truncation mutants to complement the sensitivity to 6-azauracil conferred by deletion of the *PPR2* gene, encoding TFIIS (Hubert et al., 1983; Ruet et al., 1990; Exinger and Lacroute, 1992). One additional deletion of the first 151 amino acids was constructed to provide a negative control for *in vivo* assays. It was recently reported that this deletion does not complement the phenotype of a TFIIS null (Nakanishi et al., 1995).

Subcloning Strategy

A series of amino-terminal truncations of the yeast TFIIS protein had already been constructed in pET vectors in order to express and purify them from *E. coli* for biochemical analysis (D. Awrey and A. Edwards, personal communication). I wanted the truncation proteins to be expressed at a normal physiological level to better evaluate their function relative to wild type cells. However, nothing is currently known about how production of the TFIIS protein is regulated in *Saccharomyces cerevisiae*. To avoid altering the normal expression, I constructed a TFIIS-expression plasmid containing the natural promoter and 3' sequences from the *PPR2* gene. In addition, the expression vector contains the

appropriate cloning sites to facilitate subcloning the truncated ORFs into the new expression plasmid from the parent pET vectors. The detailed strategy for constructing this yeast expression vector for the TFIIS proteins is described in the Materials and Methods (Figures 1-3).

Complementation of 6-azauracil Sensitivity *in vivo* by TFIIS Variants

The activities of the truncation variants of yeast TFIIS were assayed *in vivo* by testing the ability of each truncation to rescue sensitivity to the drug 6-azauracil disrupted for the *PPR2* gene. Deletion of the *PPR2* gene is not lethal, but cells lacking the TFIIS protein encoded by it become extremely sensitive to 6-AU (Hubert et al., 1983; Ruet et al., 1990; Exinger and Lacroute, 1992). This phenotype allowed sensitive determination of the ability of a truncation protein to restore TFIIS activity. By examining the growth of cells expressing an amino-terminally truncated TFIIS protein in place of the full length protein over a wide range of drug concentrations, it was possible to observe slight differences in the ability of the truncation mutants to rescue the drug sensitivity of *Δppr2* cells.

The first set of tests was performed on solid minimal media lacking uracil and tryptophan (SC - ura, trp) and containing various concentrations of 6-azauracil, ranging from 0 to 200 μg/ml. The absence of tryptophan ensured maintenance of the plasmid while the absence of uracil was essential since its presence rescues cells from 6-AU sensitivity (data not shown; Exinger and Lacroute, 1992). Each of the pKC16-derived expression plasmids of a TFIIS variant, pKC16(1-309), pKC16(114-309), pKC16(131-309:P292L), pKC16(144-309), and pKC16(152-309), was transformed into yeast cells (Figure 4). In addition, the plasmids pKC16 and pKJ1 were transformed into yeast as the negative and positive controls,

respectively. The pKJ1 plasmid contains the entire 3.9 kb fragment from pKC3 (Figure 2) in the shuttle vector pRS314, the original parent of the pKC16 series. Two different *Δppr2* strains were used, KJY1 and CMKY5 (Table 1), to increase confidence that any differences observed were due to the truncation proteins and not related to differences in the genetic background. It had been observed previously that CMKY5 is significantly more sensitive to 6-AU than KJY1 (K. Johnson, unpublished observations). Each set of strains (KJY1 or CMKY5 containing one of the seven plasmids) was grown to saturation in media selecting for maintenance of the plasmid. Three dilutions, 10-fold, 100-fold, and 1000-fold, of each culture were prepared and an aliquot of each was plated on each concentration of 6-AU. The growth of the cells was photographed after three days of incubation at 30°C. The growth of the KJY1 strains is shown in Figure 5, panel A and the growth of the CMKY5 strains is shown in panel B.

The abilities of the positive and negative control strains to grow in the absence or presence of 6-AU were as expected. Cells bearing the plasmid pKC16(1-309) grew indistinguishably from those bearing the positive control plasmid pKJ1. This result indicated that the mutations made in sequences flanking the ORF in order to generate the necessary cloning sites (as indicated in Figure 1) did not have a significant effect on the expression of TFIIS activity. Cells bearing the negative control plasmid pKC16, containing the *PPR2* flanking sequences but not the ORF, did not grow on media containing 6-AU. The CMKY5 strain lacking functional TFIIS was far more sensitive to the drug than the KJY1 strain lacking TFIIS. At this three day time point, 75 μg/ml 6-azauracil was required to completely inhibit the growth of the negative control strain, KJY1 + pKC16, while

25 μg/ml 6-AU was sufficient to inhibit the growth of the negative control for the other strain, CMKY5 + pKC16.

Since the comparison of growth of cells containing pKC16(1-309) with those containing pKJ1 indicated that the pKC16-derived plasmid expressed normal amounts of TFIIS activity, it was possible to examine the ability of the truncation mutants of TFIIS to complement the deletion of *PPR2*. The growth of strains containing the truncation proteins Δ2-113 and Δ2-143 looked indistinguishable from growth of the two strains containing the full length yeast TFIIS protein. These results indicate that TFIIS truncation proteins lacking the amino-terminal 143 amino acids are sufficient for wild type growth in these growth conditions.

The ability of these amino-terminal truncations of TFIIS to allow growth in the presence of 6-AU is clear evidence that this portion of the protein is not required for complementing this phenotype of the TFIIS deletion. If nuclear localization of TFIIS is regulated, the amino terminus does not play a critical role in this function. These results do not eliminate the possibility that the amino-terminus of TFIIS could be involved in negative regulation of the protein's activity or in regulation under different growth conditions.

Cells expressing the Δ2-151 TFIIS were indistinguishable from cells lacking TFIIS and did not grow on most concentrations of 6-AU (Figure 5). Thus the deletion of the additional amino acid residues from 144 and 151 eliminates the ability of an amino-terminally truncated TFIIS protein to complement deletion of the *PPR2* gene, as expected based on the previously published report (Nakanishi et al., 1995). The inability of this protein to complement will be discussed further below.

Incomplete Complementation by the Δ2-130:P292L TFIIS Protein

During construction of the deletion derivatives, the polymerase chain reaction created a mutation within one of the TFIIS ORFs. The Δ2-130 TFIIS ORF contained a single C to T change, causing the mutation of proline 292 to leucine. The plasmid expressing the Δ2-130:P292L form of TFIIS complemented the deletion of the *PPR2* gene but not as well as plasmids expressing the full length protein (Figure 5). The observance of this slight defect in both KJY1 and CMKY5 confirmed that the growth reduction was due to the form of TFIIS on the plasmid and not due a spontaneous mutation in the colonies selected from the transformation. This proline to leucine mutation is in a region of amino acids invariant amongst the cloned TFIIS sequences (Williams and Kane, 1995; also see Figure 14) and which form the zinc-ribbon domain (Qian et al., 1993a). When the aspartate and glutamate residues immediately N-terminal to this proline residue are mutated in the human protein, the resulting mutant TFIIS proteins are inactive (Ciprés-Palacín and Kane, 1995; Jeon et al., 1994). Mutation of this proline might have been predicted to be quite detrimental since proline residues often have unique structural constraints. However, TFIIS proteins containing a mutation of this proline to leucine (as described here) or to cysteine (Jeon et al., 1994) are at least partially active, indicating that this residue is not absolutely critical for the structure or the activity of TFIIS.

It is a formal possibility that the growth defect of strains expressing only the Δ2-130:P292L protein is not due to the P292L mutation but to a change in expression level or rate of degradation. This possibility would be easily tested by comparing the complementation of the Δ2-130:P292L with that of an otherwise wild type Δ2-130 protein.

To further investigate the incomplete complementation by the Δ2-130:P292L variant, the growth rate of cells containing only this form of TFIIS was determined in liquid media during exponential growth. For this experiment, four strains were examined: KJY1 cells containing pKC16, pKC16(1-309), pKC16(131-309:P292L), or pKC16(152-309). Each was inoculated into minimal media and grown overnight. The density of each culture was determined and each was diluted into two larger cultures, either minimal media or minimal media containing 50 μg/ml 6-azauracil, at an initial OD_{600} of approximately 0.01. The growth curves are shown in Figure 6 and the calculated doubling times are given in Table 3.

In the absence of 6-AU, there is no significant difference in the doubling time of the strain lacking TFIIS [KJY1 + pKC16] compared to the isogenic strain expressing wild type TFIIS [KJY1 + pKC16(1-309)]. This was the expected result based on previous measurements of KJY1 and an isogenic Ppr2+ strain (K. Johnson, unpublished data). The doubling time of the cells expressing the TFIIS variant with the P292L mutation is also the same in the absence of 6-AU as that of cells with wild type TFIIS. This is expected since the Δ2-130:P292L TFIIS protein is active, based on its ability to complement 6-AU sensitivity on solid media, and even cells without TFIIS function show no change in doubling time in the absence of 6-azauracil.

The measurement of the doubling times of the four strains in media containing 50 μg/ml 6-azauracil provides a quantitative comparison of the abilities of the Δ2-130:P292L and the full length TFIIS proteins to complement the defect conferred by deletion of the *PPR2* gene. For all four strains, the doubling times in media with 6-AU are longer than the doubling times in media without the drug, as previously observed (K. Johnson, unpublished data). The presence of this

concentration of 6-AU causes a two-fold increase in the doubling time of the strain expressing full length TFIIS. This increase is modest compared to the greater than 5-fold increase in doubling times for the strains without TFIIS or containing the non-functional Δ2-151 protein. The doubling time of the strain expressing the Δ2-130:P292L mutant TFIIS protein increased 3-fold relative to the doubling time in media lacking the drug. This is longer than that of the strain expressing wild type TFIIS protein, but is much shorter than the doubling times of both strains lacking functional TFIIS. This quantitative result agrees quite well with the more qualitative result observed with growth on solid media. This mutant protein has not yet been characterized *in vitro*, but it would be expected to be decreased in activity by those assays as well.

Western Analysis of TFIIS Expression Levels

The experiments determining ability to complement 6-AU sensitivity suggested that expression of the TFIIS proteins was sufficient for normal TFIIS function *in vivo*. However, I wanted to look directly at the level of TFIIS protein produce by expression from the reconstructed *PPR2* promoter in the pKC16-derived plasmids relative to TFIIS expression from pKJ1. A rabbit antibody raised against full length yeast TFIIS protein allowed the use of Western analysis to ascertain the protein levels of TFIIS *in vivo*. This antibody recognizes the full length protein quite well, but recognizes the truncation proteins to lesser and varying degrees (C. M. Kane, unpublished results); this is shown clearly by the controls in Figure 7.

The expression of the TFIIS proteins from the various plasmids in KJY1 was investigated. Each strain was grown in minimal media to an OD_{600} of 0.7 to 0.85.

At this point, 2 OD units of each culture were harvested and lysed in protein sample buffer. Equal amounts of each sample were resolved by SDS-PAGE, transferred to nitrocellulose and probed with the antibody against yeast TFIIS. Due to the low reactivity of the antibody with the truncation proteins, it was not possible to detect them *in vivo*. However, the full length protein was easily detected, whether expressed from pKJ1 or from pKC16(1-309). The levels of expression from the native promoter and from the reconstructed promoter were not significantly different. This provides additional confirmation that the mutations in the sequences flanking the *PPR2* ORF introduced in the construction of pKC16 do not alter TFIIS expression *in vivo*.

Summation of *in vivo* Results

The results of expression of a series of TFIIS truncation mutants in two different strains of *S. cerevisiae* demonstrate clearly that the amino-terminal portion of the protein, up to residue 143, is not needed *in vivo* for rescue of the 6-azauracil sensitivity conferred by deletion of the *PPR2* gene. Comparison of the ability of the full length protein to complement or its level of expression when produced from either pKJ1 or pKC16(1-309) provides evidence that the reconstructed *PPR2* promoter in the pKC16-derived plasmids provides normal TFIIS expression levels *in vivo*. The detection of the slight growth defect of the strain expressing the Δ2-130:P292L mutant protein indicates that the pKC16 expression vector will be a useful tool for other *in vivo* analysis of mutant TFIIS proteins.

Activities of Variant Yeast TFIIS Proteins *in vitro*

The second goal of this work was to quantitatively compare the activities of each of the truncation mutants of yeast TFIIS to stimulate stalled ternary elongation complexes containing RNA polymerase II to 1) cleave the nascent RNA transcript and 2) read-through an intrinsic block to elongation. In the previous work described in Chapter 2, I demonstrated that a truncated version of the yeast TFIIS protein lacking amino acid residues 2-113 was indistinguishable from the full length protein *in vitro*, with respect to these two activities. This work has been extended by comparing the *in vitro* activities of the Δ2-130 and the Δ2-143 proteins with those of the full length and Δ2-113 proteins (referred to as Δ1-113 in Chapter 2). Due to differences in the plasmid constructions, this Δ2-130 does not contain the P292L mutation. All four proteins were expressed as amino-terminal histidine-tagged fusions in *E. coli* and samples of each preparation are shown in Figure 8.

Stimulation of Read-through of an Intrinsic Block to Elongation

In the read-through assays described in Chapter 2, the reaction was complete before thirty minutes. Thus, for the series of *in vitro* analyses described here, timepoints were taken between five and twenty minutes after initiating the reaction. These shorter timepoints give a better indication of the kinetics of the read-through reaction.

The comparison of the abilities of the four yeast TFIIS proteins to stimulate yeast RNA polymerase II to read-through the TIa site of the human histone H3.3 gene demonstrates that all four proteins are active (Figure 9). Quantitation of read-through of the TIa site, performed in the same manner as in Chapter 2, is shown in Figure 10. To obtain accurate quantitation on the kinetics of read-through, three

assays of each variant were performed with TFIIS and RNA polymerase II at a molar ratio 8 to 1. Panels A-D of Figure 10 display assays performed with each TFIIS variant, as indicated. Panel E directly compares the four proteins; each point in this panel represents the average of either two or three of the assays shown in Panels A through D, as indicated in the figure legend. Three of the proteins, the full length, Δ2-113, and Δ2-130, stimulated read-through with nearly identical kinetics. The fourth protein, Δ2-143, stimulated read-through with slightly slower kinetics than the other three TFIIS variants.

The reduction in rate of read-through in the presence of the Δ2-143 protein (see Figure 10, Panel E) requires further comment. Although the difference in the speed of read-through stimulated by the Δ2-143 protein and by the other three proteins is very small, it is quite reproducible. A similar truncation mutant of yeast TFIIS, a Δ2-141 protein was reported to be less active in stimulating RNA polymerase II in a non-specific transcription assay (Nakanishi et al., 1995). These investigators also tested read-through, but did not report a defect in the ability of this protein to stimulate read-through. However, only a single, longer timepoint was used in their comparison, and it is likely that they missed the small difference in rate observed in the more complete timecourse reported here.

However, one additional caveat needs to be raised. Although the Δ2-143 protein appears to be slightly impaired in function *in vitro*, I have recently learned that this protein tends to aggregate and precipitate at high concentrations (D. Awrey and A. Edwards, personal communication). This property could contribute to variability in the measurement and dilution of this TFIIS variant. In light of this observation, it is possible that the read-through assays of the Δ2-143 protein were done at a lower stoichiometry than the assays of the other three TFIIS variants. If

true, this could explain the apparent reduction in rate of read-through stimulated by the Δ2-143 protein. This could be tested by comparing read-through kinetics stimulated by the Δ2-143 and the wild-type TFIIS proteins at a much larger excess of TFIIS.

Alternative Quantitation of Read-through Assays

I have been concerned for some time that the decrease in the "Percent at TIa" in the absence of TFIIS in the graphs shown in Figure 10 might represent intrinsic "escape" of stalled complexes from the TIa site. If some stalled polymerases can read-through the TIa site in the absence of TFIIS, then the amount of read-through observed in the presence of TFIIS is not completely dependent on the addition of the elongation factor. To address this possibility, the four read-through assays shown in Figure 9 were quantitated in a different manner. This alternative method of displaying quantitation of read-through assays is shown in Figure 11. For each lane, the black bar represents the intensity of the TIa band and the hatched bar represents the intensity of the Run Off band. The arbitrary unit is derived from the PhosphorImager volume integration calculation and is proportional to counts per minute (cpm). It is clear from this depiction that drop in the amount at TIa is very small in the absence of TFIIS, while the amount of R.O. transcript increases very slightly over the timecourse of the incubation without TFIIS. The combined effect is that the percent at TIa, calculated according to the formula below, decreases.

$$\% \text{ at TIa} = \frac{\text{Intensity at TIa}}{\text{Intensity at TIa + Intensity at R.O.}} * 100$$

In the absence of TFIIS, the intensity of the Run Off transcript increases more than the small decrease in the TIa signal. Possibly this slight increase is due to chasing of complexes which paused between the TIa site and the end of the transcript. Most importantly, this form of quantitation indicates that the majority of polymerases which recognize the TIa site remain stalled there in the absence of TFIIS.

This alternative method of quantitation for read-through assays also indicated that in the presence of TFIIS the amount of Run Off transcript increased more than could be accounted for merely by read-through of complexes paused at TIa. This is likely due to TFIIS-induced elongation of transcripts by complexes stalled at various locations along the template other than at the TIa site. This hypothesis is supported by the appearance of the autoradiograms; the region between TIa and the Run Off is darker in the lanes for samples containing TFIIS (lanes 3-6) than in lanes for samples lacking it (lanes 7-10). In addition, when the total intensity of each lane is determined, there is very little lane to lane variation.

Stimulation of Cleavage of the Nascent Transcript

Each variant TFIIS protein was assayed for its ability to stimulate cleavage of the nascent transcripts by RNA polymerase II ternary elongation complexes stalled at the TIa site. An example of such an assay is shown in Figure 12; this one utilized the full length protein. Each of the four TFIIS variants was assayed in the same manner and aliquots of the cleavage reaction were stopped at the same timepoints. Visually, the pattern of cleavage products and the kinetics of the appearance and persistence of each shortened RNA product was identical no matter which TFIIS was used. Quantitation, either of the percent at TIa (calculated

in the same manner as described for read-through assays) or of the intensity of the TIa band, did not indicate differences in the speed of the cleavage reaction (data not shown). Thus, there are no significant differences in the abilities of the four TFIIS variants to stimulate nascent transcript cleavage by RNA polymerase II. The activities of the Δ2-143 protein, which appears to be slightly defective in promoting read-through but not in promoting cleavage, is similar to some mutants in human TFIIS (Ciprés-Palacín and Kane, 1994).

Sequence Analysis of the Flanking Regions of the *PPR2* gene

During the construction of the pKC16 expression vector, the regions upstream and downstream of the *PPR2* ORF were sequenced. These sequences provided a basis for comparison for regions subcloned using PCR. I have found it necessary to sequence PCR-generated fragments due to the potential for errors during the polymerase chain reaction. A secondary benefit from the sequencing was the identification of the genes present on each side of the *PPR2* gene.

The results of sequencing are diagrammed in Figure 2. The sequence analysis identified a known gene on each side of the *PPR2* gene. Upstream and divergently transcribed is *RNA15* (Minvielle-Sebastia et al., 1991). However, literature reports map *RNA15* and *PPR2* to different chromosomes. *PPR2*, also called *DST1*, was mapped to chromosome VII by hybridization to a chromosome and by tetrad analysis (Clark et al., 1991). This result is consistent my results; hybridization of the 3.9 kb fragment containing *PPR2* (Figure 2) to a chromosomal blot indicated that this fragment was present on either of two chromosomes not resolved by pulse field gel electrophoresis, chromosome VII or chromosome XV (data not shown). In contrast, *RNA15* was mapped to

chromosome XVI by hybridization to a blot of yeast chromosomes (Minvielle-Sebastia et al., 1991). This discrepancy is likely to be an error in mapping rather than an additional cloning artifact in the 3.9 kb fragment containing the *PPR2* gene since part of the untranslated promoter region between the two genes is also present in the database entry for *RNA15*.

Downstream of *PPR2* and also divergently transcribed, is *HEM2*. This gene encodes δ-aminolevulinate dehydratase, the enzyme which converts δ-aminolevulinic acid to porphobilinogen, the second step in the biosynthesis of heme (Myers et al., 1987; Schlaepfer et al., 1994). The stop codons of *HEM2* and *PPR2* are separated by slightly over 1 kb of sequence.

Sequence Analysis of the Promoter Region of the *PPR2* Gene

Sequence analysis of the promoter regions of the *RNA15* and *PPR2* genes (Figure 13) identifies several potentially interesting features. Although further analysis will be required to determine their significance, they provide a basis for speculation into the regulation of the *PPR2* gene. The ATG codons of the two ORFs are separated by 446 base pairs. There are at least two sequences in this region that match the TATA box consensus sequence (reviewed in Struhl, 1989). The first one is close to *RNA15* and is oriented in the appropriate direction for transcription of that gene. The second one is almost in the middle, slightly closer to *PPR2* than to *RNA15*, and is identical to the TATA box of the *GAL1* gene (Selleck and Majors, 1987).

Intriguingly, there are sequences which match the consensus binding sites for three yeast proteins involved in transcriptional regulation. There are three putative sites for *GCN4*, one close to *RNA15* and two close to *PPR2*. One of these,

labelled site 1 (Figure 13), is a perfect match for the *GCN4* consensus core binding sequence 5'-TGACTC-3' (Hope and Struhl, 1985; Hill et al., 1986). The other two differ from the consensus by one base each. The *GCN4* protein has been shown to bind to sites with either of these differences from the consensus, although with lower affinity (Arndt and Fink, 1986). Almost equidistant between the two ORFs, is a putative *ABF1* binding site (Buchman and Kornberg, 1990; McBroom and Sadowski, 1994a, 1994b). Slightly closer to the *PPR2* gene, is a perfect match to the consensus binding sequence (5'-YNNYYACCCG-3') for the *GRF2* protein. The site is in the appropriate orientation for respect to transcription of *PPR2* (Chasman et al. 1990). Whether these proteins bind to these sites or are involved in the regulation of either gene remains to be determined.

DISCUSSION

The results presented in this work as well as those of Nakanishi et al. (1995), indicate clearly that the carboxyl terminus of TFIIS is sufficient for the known activities of the protein both *in vitro* and *in vivo*. I have demonstrated here that deletion of the first 143 amino acids does not impair the ability of TFIIS to stimulate RNA polymerase II ternary elongation complexes to cleave the nascent transcript, to read through an intrinsic block to elongation, or to rescue the sensitivity of *Δppr2* cells to the drug 6-azauracil. Although TFIIS has been characterized extensively *in vitro*, the function of TFIIS in cells remains unknown. Based on its ubiquitous presence in eukaryotic cells and fairly high degree of sequence conservation (Williams and Kane, 1995), it seems likely that TFIIS has important functions related to the process of polymerizing RNA within cells.

Although nothing is currently known about the regulation of TFIIS expression in *S. cerevisiae*, the sequence analysis described here suggests some interesting possibilities. Putative binding sites for three yeast DNA binding proteins are present in the promoter region between the *RNA15* and *PPR2* ORFs. Two of these factors, ABF1 and GRF2 (also called REB1, Factor Y, or QBP), are abundant proteins with multiple functions. ABF1 is involved in transcriptional activation, transcriptional silencing of *HML* and *HMR*, and DNA replication (Buchman et al., 1988; Buchman and Kornberg, 1990) and bends DNA when it binds (McBroom and Sadowski, 1994a, 1994b). It is thought that a function of ABF1 in these various roles may be to remodel chromatin structure. Likewise, the GRF2 protein is implicated in modulation of chromatin structure (Chasman et al., 1990; Erkine et al., 1995). Alone GRF2 stimulates only modest levels of transcription, but when

combined with other activators it contributes to high levels of activation (Remacle and Holmberg, 1992).

The third protein with putative binding sites within the *PPR2* promoter region is GCN4, a transcription factor required for the general amino acid control derepression response, where the relatively high basal expression levels of many amino acid biosynthetic genes are increased in response to starvation for any single amino acid (Jones and Fink, 1982). The presence of putative GCN4 binding sites suggests that *PPR2* transcription could be increased in response to amino acid starvation. Perhaps the increase in transcription of the genes encoding biosynthetic enzymes is facilitated by an increase in the level of the TFIIS transcription elongation factor. It will be very interesting to investigate the possible involvement of these regulatory factors in the expression of the *PPR2* gene.

As the carboxyl terminus of TFIIS is sufficient for the known activities of the protein both *in vitro* and *in vivo*, analysis of mutations in this region may begin to elucidate how the protein interacts with RNA polymerase II. Although it has been proposed that TFIIS contains an RNA polymerase binding domain in the middle of the protein and a nucleic acid binding domain in the carboxyl-terminus (Agarwal et al., 1991; Qian et al., 1993a, 1993b), the evidence for these domains is unconvincing. Since studies of transcript cleavage stimulated by TFIIS or by pyrophosphate suggest that the polymerase active site catalyzes the transcript cleavage reaction (Rudd et al., 1994; R. Weilbaecher, unpublished results), it may be that TFIIS does not interact with nucleic acid at all. Yet it is clear that TFIIS does interact with RNA polymerase II and that the carboxyl-terminus is sufficient for this. The existing data do not distinguish between residues important for the structure and those which interact with the polymerase. Further mutagenesis and

structural analysis of the TFIIS protein will elucidate how this factor interacts with RNA polymerase to promote the ubiquitously conserved transcript cleavage reaction.

Previously published alignments of TFIIS proteins from human and *S. cerevisiae* placed a gap in the human protein such that the four residues of the yeast protein indicated by the single underline in Figure 14 were not aligned with the other TFIIS proteins (Williams and Kane, 1995; Chen et al., 1992). In that alignment, a portion of which is shown in Figure 14, Panel A, the first arginine in the region spanned by mutant 1 of the human protein is aligned with histidine 145 in the *S. cerevisiae* protein. In the alignment of Williams and Kane, all other TFIIS proteins have an arginine residue in this position. Consideration of the mutant TFIIS proteins reported here as well as those reported by Nakanishi et al. (1995) and by Ciprés-Palacín and Kane (1994, 1995) suggests a possible shift in the alignment of the *S. cerevisiae* TFIIS protein with respect to the other reported sequences.

Panel B of Figure 14 realigns the shortest fragments of the yeast and human TFIIS proteins known to be functional (Nakanishi et al., 1995; Ciprés-Palacín and Kane, 1994, 1995). It is striking that the majority of mutations made in the carboxyl-terminus by alanine scanning, a technique intended to minimize structural perturbations, have little effect on the activity of TFIIS (Ciprés-Palacín and Kane, 1995). In the completely inactive human TFIIS mutant 1 of Ciprés-Palacín and Kane (indicated by the dotted underline), four charged residues within a six residue span, including the conserved arginine, were changed to alanines. The yeast Δ3-147 TFIIS protein reported by Nakanishi et al. is active, but deletion of the next two residues, leucine and arginine, produces an inactive protein. Since

the deletion of histidine 145 is inconsequential to the activity of the yeast TFIIS protein, but deletion of leucine 148 and arginine 149 is not, this arginine residue seems more comparable to the first arginine residue within the span of human mutant 1. By realigning the *S. cerevisiae* sequence, yeast arginine 149 no longer aligns with human arginine 145. However this second arginine within the mutant 1 region is not invariant amongst all the mammalian TFIIS sequences. The new alignment is more consistent with the currently available data on residues important for TFIIS function. In it, the arginine residue in the mutant 1 region is invariant and the residue immediately amino-terminal is conserved. It would be very interesting to make single point mutations in each amino acid changed as a cassette in mutant 1. Based on the activities of known mutant TFIIS proteins and this new alignment, the change in the first arginine residue would be predicted to be the major cause of the inactivity of mutant 1.

The difference in activity of the Δ3-147 and Δ2-149 truncation proteins certainly suggests an essential function for leucine 148 and arginine 149, but the nature of this function is not clear. One or both of these residues may be required for proper structure of the TFIIS carboxyl-terminus. Alternatively, these residues may interact with the polymerase to promote transcript cleavage. Based on the evidence that the catalytic site of the polymerase is responsible for catalyzing the transcript cleavage reaction as well, it is hypothesized that the function of TFIIS is to reposition the catalytic site of stalled ternary complexes. In this model, TFIIS might stimulate or stabilize a particular conformation of the polymerase (Gu et al., 1993; Reines, 1992, 1994; Chamberlin, 1994). Mutant TFIIS proteins which stimulate transcript cleavage but not read-through suggest that TFIIS may be required for maintenance of this conformation as well (Ciprés-Palacín and Kane, 1994, 1995)

Many questions remain about the function of TFIIS *in vivo*. The function of the amino-terminus of the protein remains obscure. The ability of the carboxyl-terminal portion of TFIIS to fully complement the 6-AU sensitivity of cells lacking the *PPR2* gene indicates that the amino-terminus is not required for localization or activation of TFIIS activity in the tested growth conditions. A possible function in negative regulation of TFIIS activity by the amino-terminus is not excluded by this result. Although there is evidence for tissue-specific expression of various TFIIS genes in humans and mice (Kanai et al., 1991; Xu et al., 1994; Weaver, 1995), it is not even certain that TFIIS activity is regulated in *S. cerevisiae*. The TFIIS expression vector described here will be a useful tool to investigate the effects of mutants in the TFIIS protein and the regulation of the expression of TFIIS activity.

TABLE LEGENDS

Table 1 - List of Strains

The names and relevant genotypes of strains of *Escherichia coli* and *Saccharomyces cerevisiae* are given. The *E. coli* strains DH10B (Grant et al., 1990) and XL-1Blue (Bullock et al., 1987) were obtained from Stratagene. *S. cerevisiae* strains were obtained from several sources as follows. YPH499 (Sikorski and Hieter, 1989) was obtained from J. Thorner. CH1305 was received from C. Holm (Kranz and Holm, 1990). KJY1 and CMKY5 were constructed by K. Johnson (unpublished results). CMKY3 and CMKY4 were constructed by J. Davie (unpublished results). The hisG-*URA3*-hisG cassette is as described by Alani et al. (1987).

Table 1 - List of Strains

Escherichia coli	Genotype
DH10B	F⁻ *araD139* Δ(*ara, leu*)7697 Δ*lacX74 galU galK rpsL deoR* Φ80d*lacZ*ΔM15 *endA1 nupG recA1 mcrA* Δ(*mrr hsdRMS mcrBC*)
XL-1Blue	*recA1 endA1 gyrA96 thi hsdRIT*($r^-_K m^+_K$) *SupE44 relA1* λ^- (*lac*) {*F' proAB lac1*^{q} *ZΔM15 Tn10*(*tet*^{R}})

Saccharomyces cerevisiae	Genotype
YPH499	MAT**a** ura3-52 lys2-801^{am} ade2-101^{oc} trp1-Δ63 his3-Δ200 leu2-Δ1
KJY1	YPH499 Δppr2::hisG-URA3-hisG
CH1305	MAT**a** ade2 ade3 leu2 ura3 lys2
CMKY3	CH1305 Δppr2::hisG-URA3-hisG
CMKY4	CMKY3 Δppr2::hisG
CMKY5	CMKY4 Δtrp1::hisG-URA3-hisG

Table 2 - Sequencing Primers for Region Containing *PPR2*

The sequence of each primer is given in the 5'-->3' direction. A 'D' at the end of the name of the primer indicates a primer in the downstream direction with respect to transcription of the *PPR2* gene. A 'U' at the end of the primer name indicates a primer in the upstream direction. The general location of a primer is indicated as being within the open reading frame (ORF) of *PPR2*, upstream of the ORF (5'), or downstream of the ORF (3'). T3+ and T7+ primers anneal to sequences in the pBS.KS+ vector regions flanking the insert in pKC3.
As pRS314 is derived from pBS.KS+, these primers are also present and on the same sides of the *PPR2* gene in pKC16 and its derivatives.

Table 2 - Sequencing Primers for Region Containing *PPR2*

Primer Name	Complementary Region	Location
T3+	GCAATTAACCCTCACTAAAGGG	5'
T7+	GCGTAATACGACTCACTATAGG	3'
P37-1D	GCCATGAAACAAGACAAG	ORF
P37-2U	CTAAGAGAATTTCCATCTG	ORF
P37-5D	CTCGTGAAGAAAATGATTAGC	ORF
P37-7U	GGGCGCTGGTGCCAGATCC	ORF
P37-8U	TAAGATTTCTAGAACTGCAGC	ORF
P37-9D	CAATTGCAAACAAGATCTGCG	ORF
P37-10U	CATTTCACTTTCTATGGCC	ORF
P37-11U	CCTGGCTTTGTAAGCGGC	ORF
P37-12D	GTCATATCAAAGAATAACCC	ORF
P37-16D	CTATGCATTAAATATGGCTCGG	3'
P37-17U	GTTAGAGGGATATACAAGAAAGG	5'
P37-18D	CCTATGCAGCTAAGTTTAGCGG	3'
P37-19D	TGGAACGTCGATACCAGAGGG	5'
P37-20U	CGGGAGTTTCACAACAGCACC	5'
P37-21D	CGACCTACCAGTTTGGGGGTCG	5'
P37-22D	GCAAGTGCTAGAAGTGACTCGG	5'
P37-23D	CAGCATCTTCAGGATGCGCTC	5'
P37-24U	GGTAGGTCGAAAGGGTACGCG	5'
P37-25U	CGAGTGACTCTGGAACTATGC	3'
P37-26D	CAGGCCACGACCGGCAGGAGG	3'
P37-27D	CATAAAGGTGCAGAAACTCATGG	3'
P37-28U	GCGTGTCTCAGATCTCTGTCC	3'
P37-29D	CCTATCTGTCTTACTAGCGCC	3'
P37-32U	TTGGATATGGTTGTATGTTTGGC	3'
P37-33U	GGTTTCTTGAGGGCGGGTGCC	3'
P37-34D	CTCTTACCCGCCCCACTGTGC	5'

Table 3 - Doubling Times of KJY1 Strains With or Without 6-azauracil

The doubling times during logarithmic growth in either SC-trp, ura or SC-trp, ura with 50 μg/ml 6-azauracil were determined for the strains KJY1 + pKC16, KJY1 + pKC16(1-309), KJY1 + pKC16(131-309:P292L), KJY1 + pKC16(152-309). Based on the semi-logarithmic plots shown in Figure 6, the time spent in exponential growth was determined. Doubling times were calculated according to the formula:

$$t\,½ = \frac{(Total\ Time) * \ln 2}{\ln(final\ OD / initial\ OD)}$$

Table 3 - Doubling Times of KJY1 Strains in Media With or Without 6-azauracil

Media	SC - trp, ura	SC - trp, ura with 50 ug/ml 6-azauracil
Plasmid in KJY1	Doubling Time (hours)	
pKC16	1.7	10.2
pKC16(1-309)	1.6	3.7
pKC16(131-309:P292L)	1.7	5.5
pKC16(152-309)	1.9	10.1

FIGURE LEGENDS

Figure 1 - Polylinker of pKC-RS314

The sequence of one strand of the polylinker of pKC-RS314 (KRC-POLY) is given, from the KpnI site to the SacI site. The cleavage site for each enzyme is indicated with an arrow. The box indicates sequence immediately upstream of the *PPR2* ORF included in the polylinker sequence since the chosen cloning scheme did not include this sequence as part of either the 5'-fragment or the 3'-fragment. Positions which will produce base changes from the natural sequence of the *PPR2* gene are underlined once. A three base insertion, relative to the *PPR2* sequence, is indicated with the triangle. The double underline indicates sequence which will be the ATG codon of the variant yeast TFIIS proteins when cloned in between the NdeI and BamHI sites. A site for XbaI is indicated but will not cut when the DNA is methylated by a dam^+ strain.

Polylinker of pKCRS314

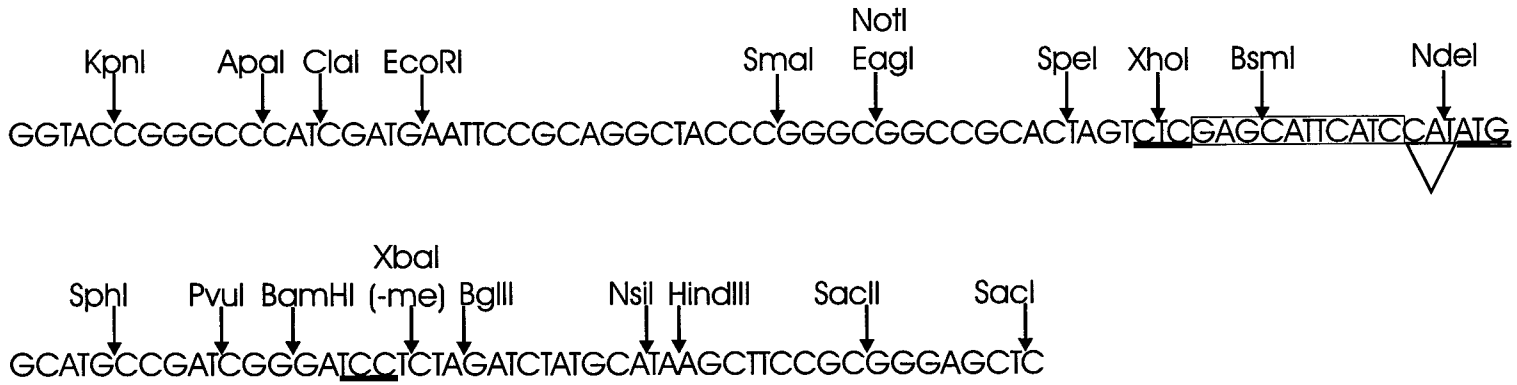
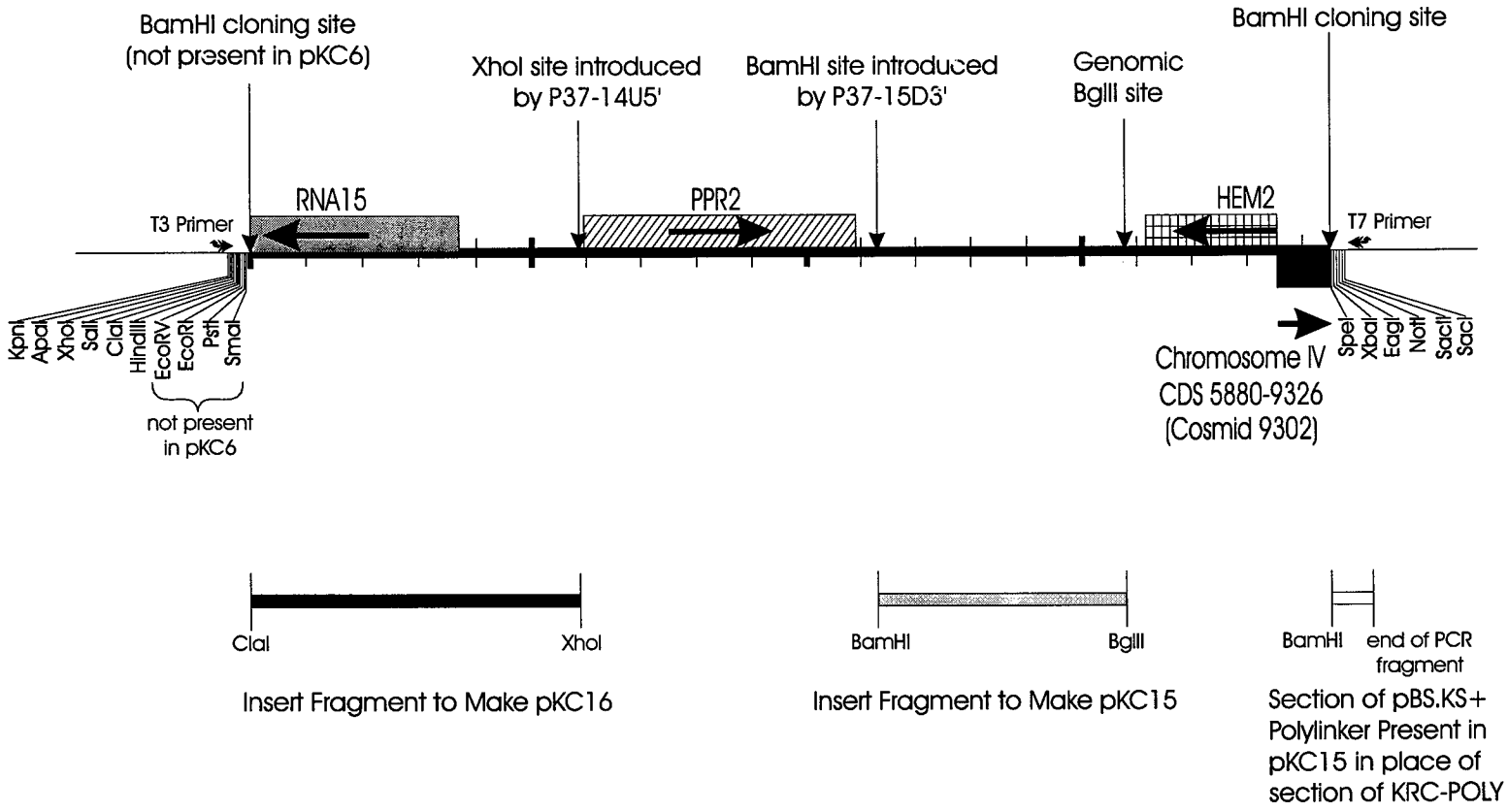


Figure 2 - Insert of pKC3/pKC6 and Subcloning Strategies for pKC15 and pKC16

The 3.9 kb BamHI fragment containing the *PPR2* gene (subcloned from pFL44D into pBS.KS+ to make pKC3; described in Chapter 2) was sequenced in its entirety and is diagrammed with the polylinker in pKC3. In addition to *PPR2*, two partial open reading frames were found. *RNA15* is upstream of *PPR2* and is divergently transcribed. *HEM2* is downstream of *PPR2*. The position and direction of each is indicated. In addition, adjacent to the fragment of *HEM2*, is an approximately 200 base pair piece of chromosome IV. This DNA is a section from the middle of a large CDS of unknown identity and function. The junction between the *HEM2* sequence and the chromosome IV DNA is an MboI site (a four-cutting enzyme which produces ends compatible with BamHI). Presumably, this short fragment was ligated to the fragment containing *PPR2* during the original genomic library construction. The plasmid pKC6 is identical to pKC3, with the exception of the removal of several polylinker sites, most notably the BamHI site. This plasmid was used as the template for the PCR-based cloning strategy of the 5'-fragment since it lacks the BamHI site on the upstream side of *PPR2*. The position of restriction sites which were added using PCR primers are indicated by arrows and primers annealing to the pKC6 polylinker are also indicated. The 3'- and 5'-fragments for subcloning are diagrammed below with the appropriate restriction sites. In addition, a polylinker fragment is depicted since it was fortuitously subcloned into pKC-RS314 along with the 3'-fragment to make pKC15 and replaces the KRC-POLY sequence. The 5'-fragment was subcloned into pKC15 to make pKC16.

Insert of pKC3/pKC6 and Cloning Strategy for pKC15 and pKC16

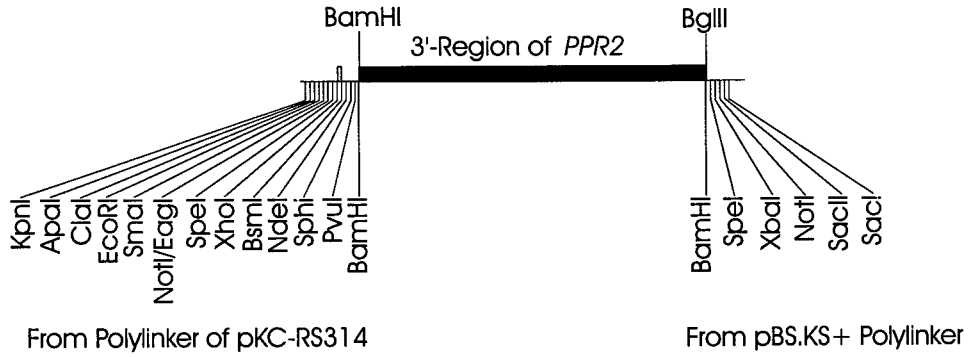


Scale : 7.5 mm = 200 base pairs

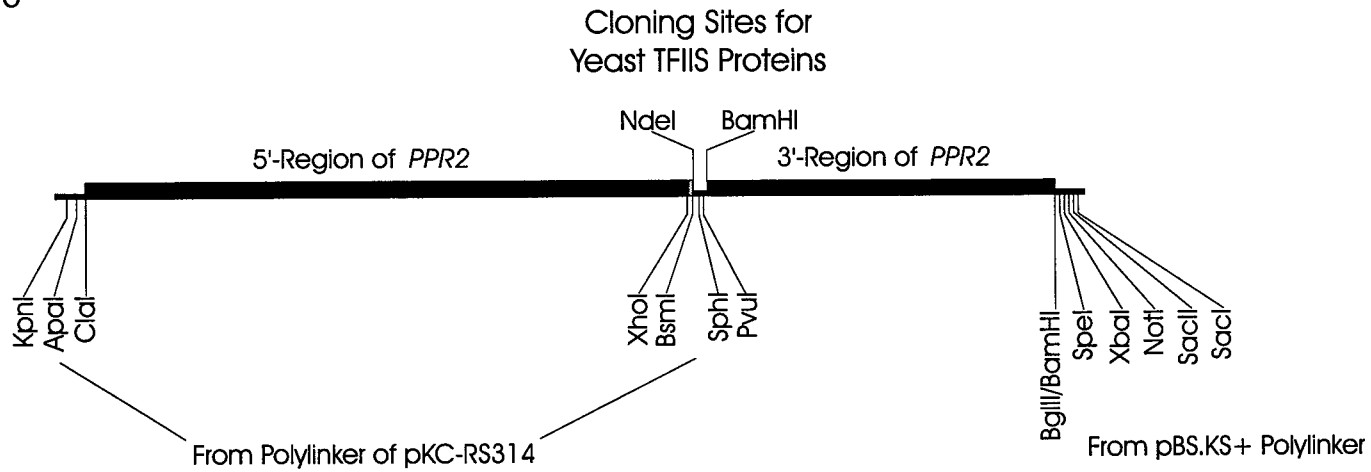
Figure 3 - Inserts and Relevant Sites of pKC15 and pKC16

A. pKC15. The 3'-Region downstream of the *PPR2* gene was subcloned into pKC-RS314 to make pKC15. The 3'-region is depicted by the black rectangle. The minuscule light gray rectangle indicates a small region of the *PPR2* gene which is encoded by the polylinker of pKC-RS314. **B. pKC16**. The 5'-region upstream of the *PPR2* gene, including the upstream fragment of the *RNA15* ORF, was subcloned into pKC15 to make pKC16. The 5'-region is depicted by the dark gray rectangle. The black and light gray rectangles are as those described for pKC15. The NdeI and BamHI sites are to be used for insertion of various yeast TFIIS open reading frames, subcloned out of pET11a or pET15b.

A pKC15



B pKC16



Scale : 1 cm = 150 base pairs

Figure 4 - Activities of Variant Yeast TFIIS Proteins

The full length and truncated yeast TFIIS open reading frames are depicted by gray rectangles. For both the *in vitro* and *in vivo* assays, a '+' indicates the protein was active, a '-' indicates the protein was not active, and ND indicates the protein was not tested. All of the *in vitro* assays were performed with yeast TFIIS proteins containing amino-terminal fusions of a short peptide containing a run of six histidine residues, not depicted here. The histidine containing tag was not present in the proteins expressed *in vivo*. For the *in vivo* assays, the Δ2-130 protein contained an additional point mutation, changing proline 292 to leucine. This protein complemented sensitivity to 6-azauracil slightly less well than the full length, but grew much better than cells with no or inactive TFIIS; therefore it is marked with a '+'.

Activities of Variant Yeast TFIIS Proteins

Yeast TFIIS Protein		Active *in vitro* for Read-through ana Cleavage[a]	Active *in vivo* for Rescue of 6-AU Sensitivity
1-309		+	+
Δ2-114		+	+
Δ2-130		+	+[b]
Δ2-143		+	+
Δ2-151		N.D.	-

(a) All proteins used for *in vitro* experiments were fused to a Histidine tag

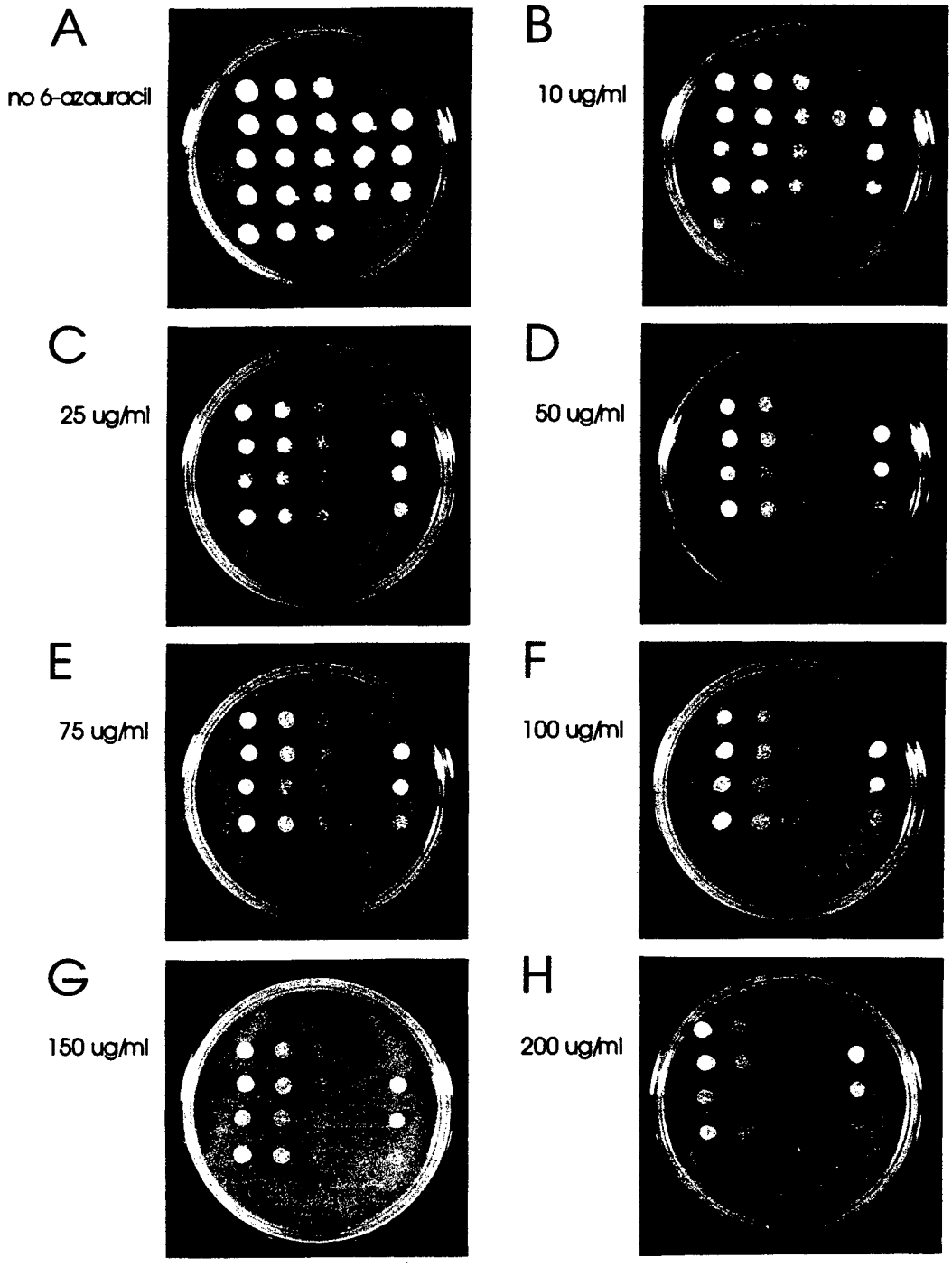
(b) Tested with the mutant Δ2-130:P292L

Figure 5 - Complementation of 6-azauracil Sensitivity by TFIIS Variants

Cells of each strain containing one of the following plasmids: pKC16(1-309), pKC16(114-309), pKC16(131-309:P292L), pKC16(144-309), pKC16(152-309), pKC16, or pKJ1 were replica plated (as described in the Materials and Methods) onto media lacking or containing 6-azauracil at the concentration indicated to the left of each panel. The plates were photographed after 3 days growth at 30°C. The plasmid present and the fold-dilution placed in each well is diagrammed in the following chart. **A.** Growth of the KJY1 strains is shown. **B.** Growth of the CMKY5 strains is shown.

Plasmid				pKC16	pKJ1
pKC16(1-309)	10x	100x	1000x	-	-
pKC16(114-309)	10x	100x	1000x	10x	10x
pKC16(131-309:P292L)	10x	100x	1000x	100x	100x
pKC16(144-309)	10x	100x	1000x	1000x	1000x
pKC16(152-309)	10x	100x	1000x	-	-

Growth of KJY1 Strains



Growth of CMKY5 Strains

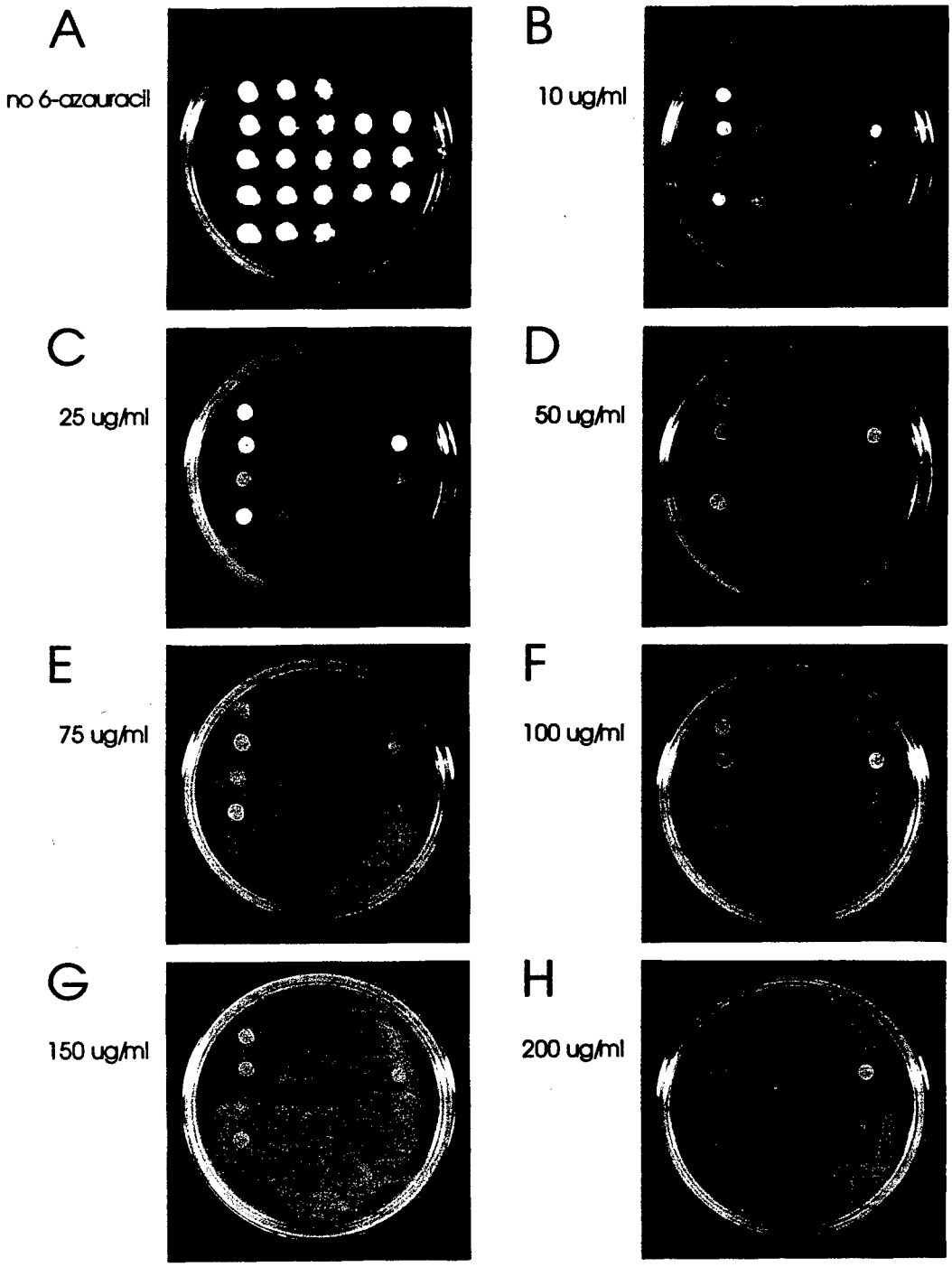


Figure 6 - Growth Curves of KJY1 Strains With or Without 6-azauracil

Small cultures of KJY1 containing one of the plasmids pKC16, pKC16(1-309), pKC16(131-309):P292L, or pKC16(152-309) were grown overnight at 30°C to allow each strain to begin growing vigorously. Each culture was used to inoculate two cultures (one with and one without 6-AU) to an OD_{600} of approximately 0.01. The cultures were shaken at 30°C and aliquots were removed at various times to determine the optical density. OD_{600} is plotted with respect to time. **A.** Growth of the four cultures in SC-trp, ura is plotted. **B.** Growth of the four cultures in SC-trp, ura + 50 μg/ml 6-azauracil is plotted. Note that the scales of both the x and the y axes are not the same for panels A. and B.

Growth of KJY1 Strains

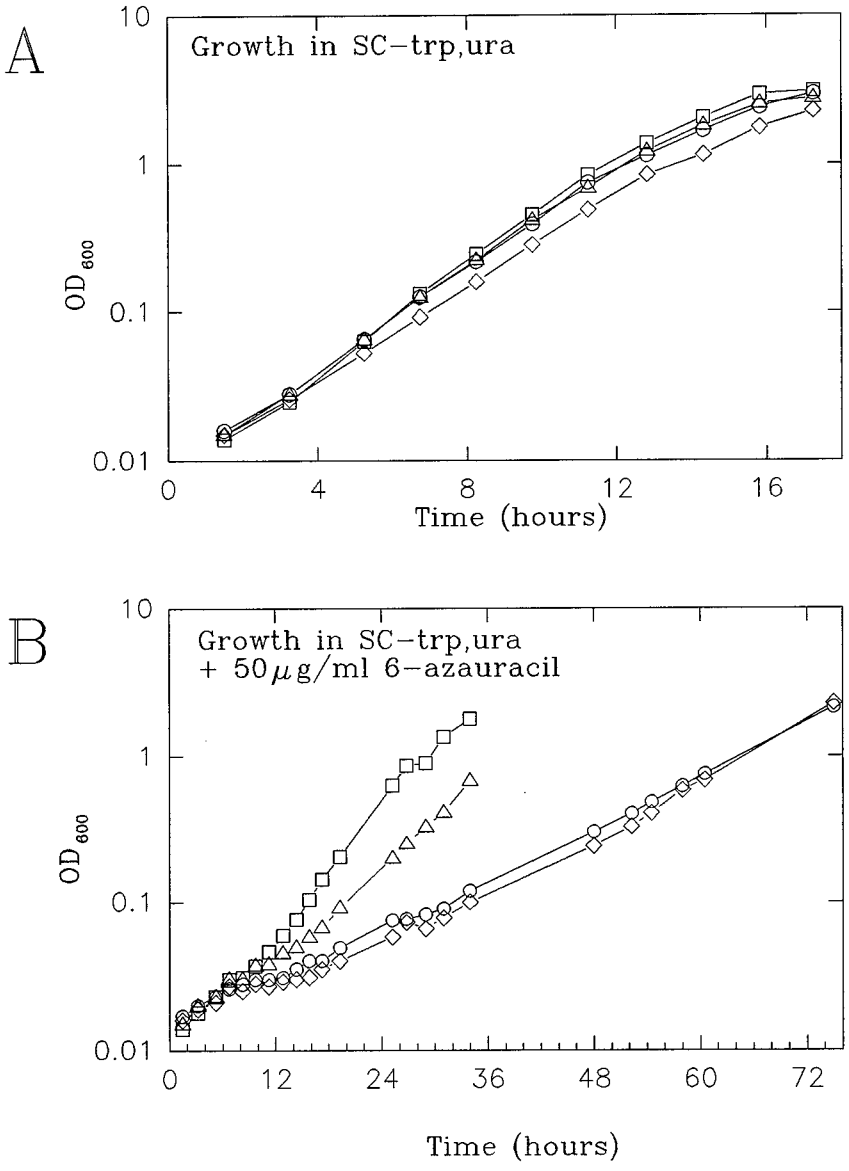


Figure 7 - Western to Detect TFIIS Expression in KJY1 Strains Bearing Plasmids

Samples of purified yeast TFIIS proteins or cell extracts from KJY1 strains bearing plasmids expressing a TFIIS variant or a control plasmid were resolved on a 15% SDS polyacrylamide gel. In lanes 1-4 are 100 ng samples of purified His-yTFIIS(1-309), His-yTFIIS(Δ2-113), His-yTFIIS(Δ2-130), or His-yTFIIS(Δ2-143), respectively. In lanes 5-11 are samples of KJY1 strains bearing the plasmid indicated.
Protein extracts were prepared from plasmid-bearing cells grown to similar optical densities in mid to late exponential growth. Equivalent amounts of each of these samples was loaded in each lane. The proteins were transferred onto a nitrocellulose membrane and probed with antibodies for chemiluminescent detection. The primary antibody was a rabbit IgG raised against full length yeast TFIIS protein. The secondary antibody was donkey anti-rabbit IgG coupled to horseradish peroxidase. Chemiluminescence detection was performed as described in the Materials and Methods. This is the 5' exposure.

Control Proteins				KJY1 Strains						
His-yTFIIS(1-309)	His-yTFIIS(Δ1-113)	His-yTFIIS(Δ1-130)	His-yTFIIS(Δ1-143)	KJY1 + pKJ1	KJY1 + pKC16	KJY1 + pKC16(1-309)	KJY1 + pKC16(114-309)	KJY1 + pKC16(131-309)	KJY1 + pKC16(144-309)	KJY1 + pKC16(152-309)
1	2	3	4	5	6	7	8	9	10	11

Figure 8 - Preparations of Histidine-tagged Yeast TFIIS Proteins

Samples of purified yeast TFIIS proteins were resolved on a 15% SDS polyacrylamide gel and visualized by silver staining. Lanes 1-4 contain His-yTFIIS(1-309), His-yTFIIS(Δ2-113), His-yTFIIS(Δ2-130), and His-yTFIIS(Δ2-143), respectively. Marker lane contains Silver Stain Low Molecular Weight Protein Standards from BioRad, sizes in kDa as indicated. **A.** Samples of TFIIS proteins are 250 ng each. **B.** Samples of TFIIS proteins are 500 ng each. BSA is also present in the sample in lane 4.

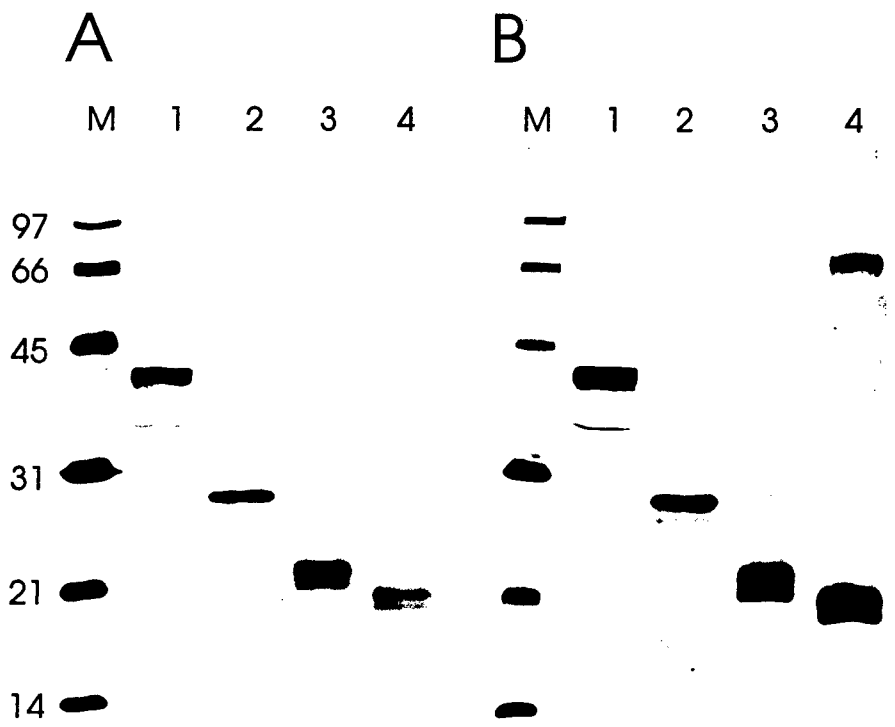

A
B
M 1 2 3 4
M 1 2 3 4
97
66
45
31
21
14

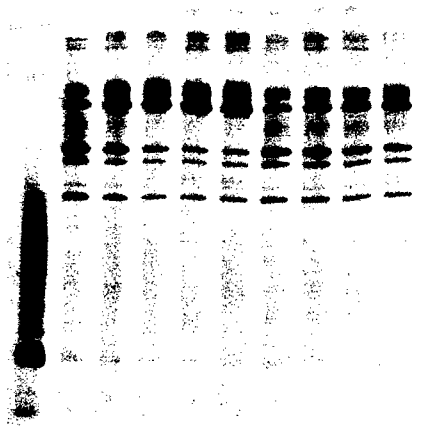
Figure 9 - Read-through Stimulated by Variant Yeast TFIIS Proteins

Ternary complexes containing RNA polymerase IIa were formed as described in Materials and Methods. Aliquots of each were stopped at various times. Lane 1: RNA from initiated complexes stopped at 1 minute; Lane 2: RNA from ternary complexes chased to the TIa, TIb, and TII sites or to the run off transcript, stopped at 2 minutes. At 2 minutes 30 seconds, 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. Lanes 3-6: RNA from ternary complexes incubated with TFIIS, stopped after 5, 10, 15, or 20 minutes of incubation, respectively; Lanes 7-10: RNA from ternary complexes incubated in the absence of TFIIS, stopped after 5, 10, 15, or 20 minutes of incubation, respectively. The molar ratio of TFIIS to RNA polymerase II was approximately 8:1 for the reactions shown in the four panels. **A.** Reactions containing His-yTFIIS(1-309). **B.** Reactions containing His-yTFIIS(114-309). **C.** Reactions containing His-yTFIIS(131-309). **D.** Reactions containing His-yTFIIS(144-309).

A

His-yTFIIS(1-309)

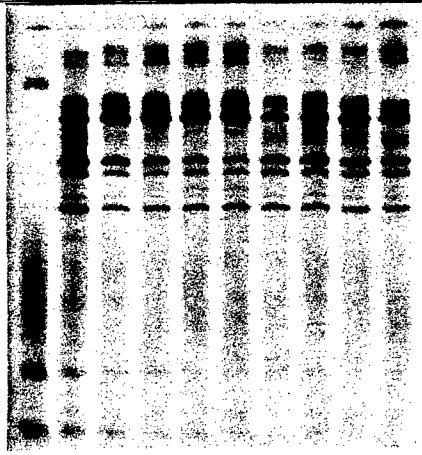
Lane	1	2	3	4	5	6	7	8	9	10
TFIIS	–	–	+	+	+	+	–	–	–	–
Time	1′	2′	5′	10′	15′	20′	5′	10′	15′	20′



B

His-yTFIIS(114-309)

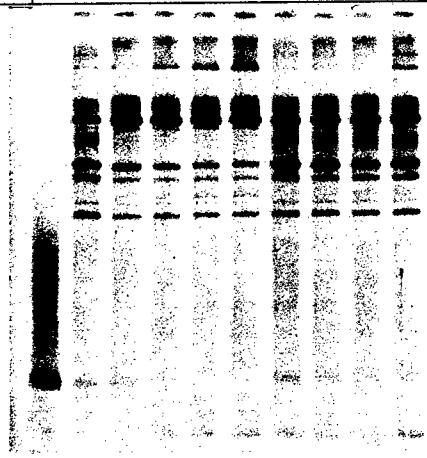
Lane	1	2	3	4	5	6	7	8	9	10
TFIIS	–	–	+	+	+	+	–	–	–	–
Time	1′	2′	5′	10′	15′	20′	5′	10′	15′	20′



C

His-yTFIIS(131-309)

Lane	1	2	3	4	5	6	7	8	9	10
TFIIS	–	–	+	+	+	+	–	–	–	–
Time	1′	2′	5′	10′	15′	20′	5′	10′	15′	20′



D

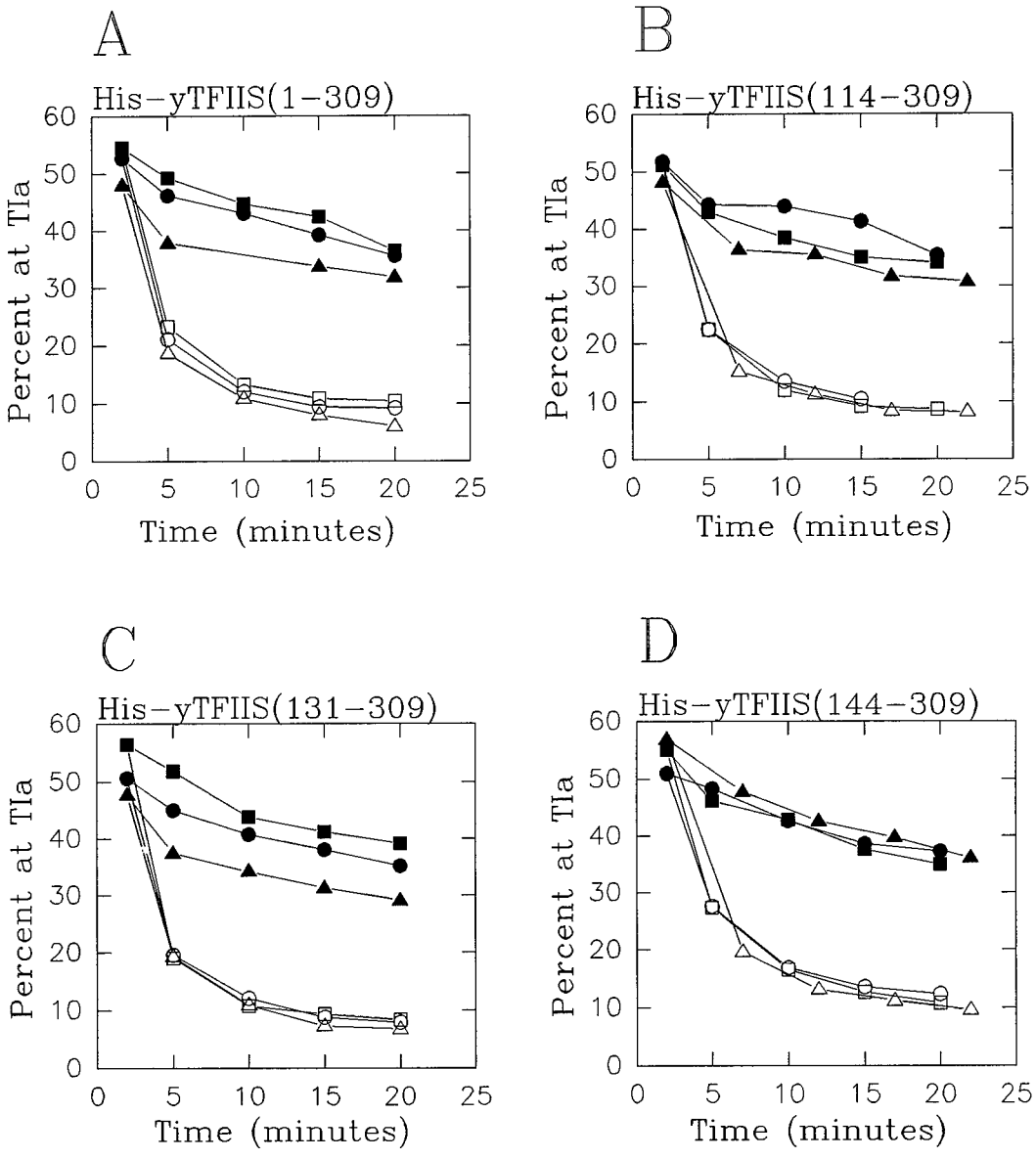
His-yTFIIS(144-309)

Lane	1	2	3	4	5	6	7	8	9	10
TFIIS	–	–	+	+	+	+	–	–	–	–
Time	1′	2′	5′	10′	15′	20′	5′	10′	15′	20′

Figure 10 - Quantitation of Read-through Stimulated by Variant TFIIS Proteins

Elongation assays with RNA polymerase IIa were performed as described in Chapter 2. One of the variant yeast TFIIS proteins was added and reactions were incubated for the times indicated. RNAs were quantitated as described in Materials and Methods. Panels A-D display the results of three separate assays of each protein. Each assay is represented with a different symbol. Filled symbols represent timepoints in the absence of TFIIS; open symbols represent timepoints in the presence of TFIIS. The *y*-axis represents the percent of RNAs with 3'-ends at the TIa site in the absence of presence of a TFIIS protein at an RNA polymerase II to TFIIS ration of 8:1. **A.** Reactions containing His-yTFIIS(1-309). **B.** Reactions containing His-yTFIIS(114-309). **C.** Reactions containing His-yTFIIS(131-309). **D.** Reactions containing His-yTFIIS(144-309). **E.** Comparison of each of the four variant yeast TFIIS proteins. Each point represents the average of three experiments [for His-yTFIIS(1-309) and His-yTFIIS(131-309)] or two experiments [for His-yTFIIS(114-309) and His-yTFIIS(144-309)].

Quantitation of Read-through Stimulated
by Variant Yeast TFIIS Proteins



Quantitation of Read-through Stimulated
by Variant TFIIS Proteins

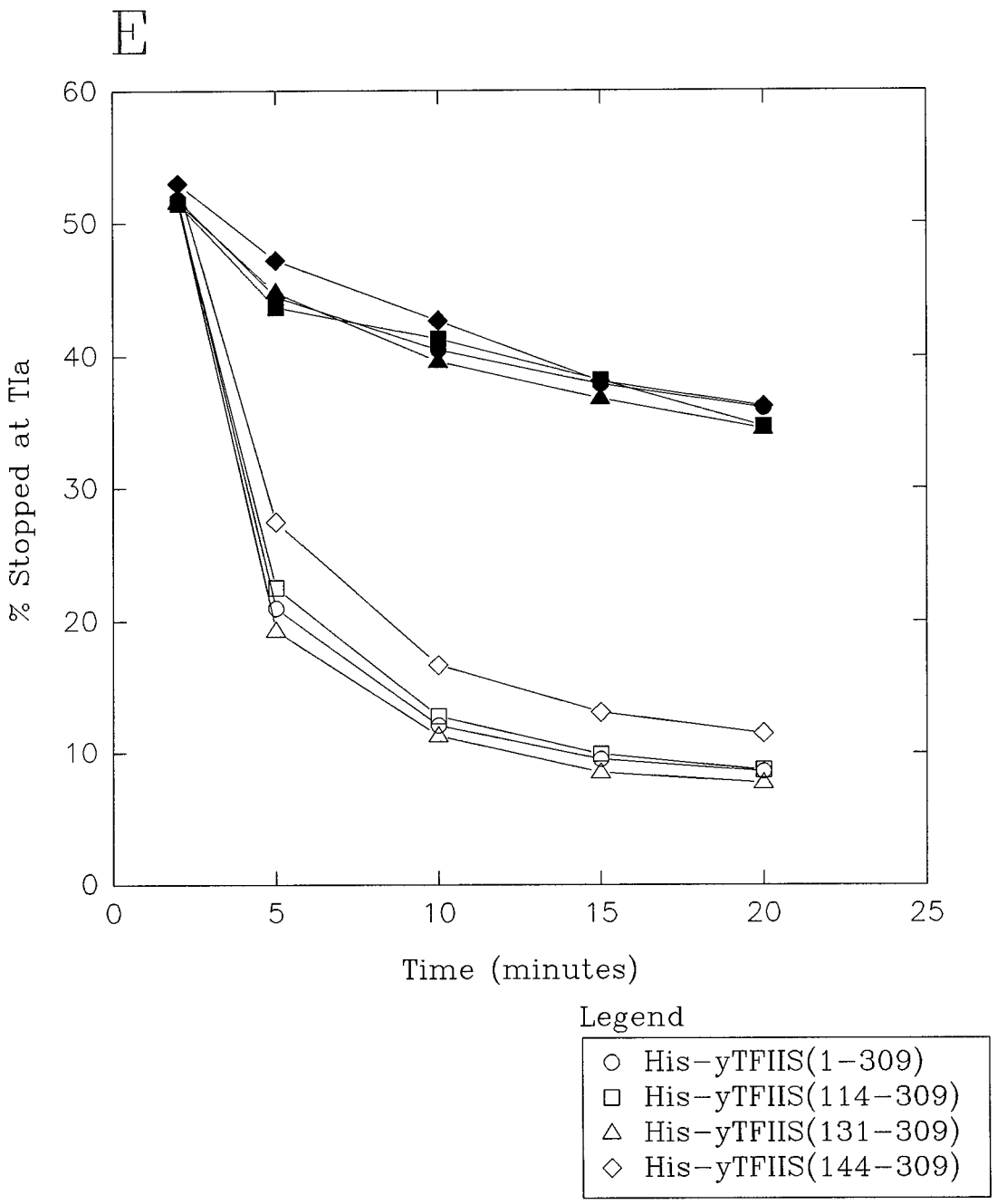


Figure 11 - Intensities of the Run-off and TIa Bands in Read-through Assays

The intensities determined (using Molecular Dynamics ImageQuant software) for the Run-off and TIa bands in each of the gels shown in Figure 9 are plotted here. The filled bars indicate the intensity of the TIa band in each lane; the hatched bars indicate the intensity of the Run-off band in each lane. In each panel, lanes 3-6 contain TFIIS and lanes 7-10 do not. See the legend to Figure 9 for more details. (Note: Panels A, B, and D represent three assays performed on the same day with the variant of histidine-tagged yeast TFIIS indicated above each panel. Panel C represents an assay from a set performed four days earlier with the same batch of [α^{32}P]-CTP. This affects only the scale of the *y*-axis.)

Relative Intensities of RunOff and TIa Bands In Read-through Assays

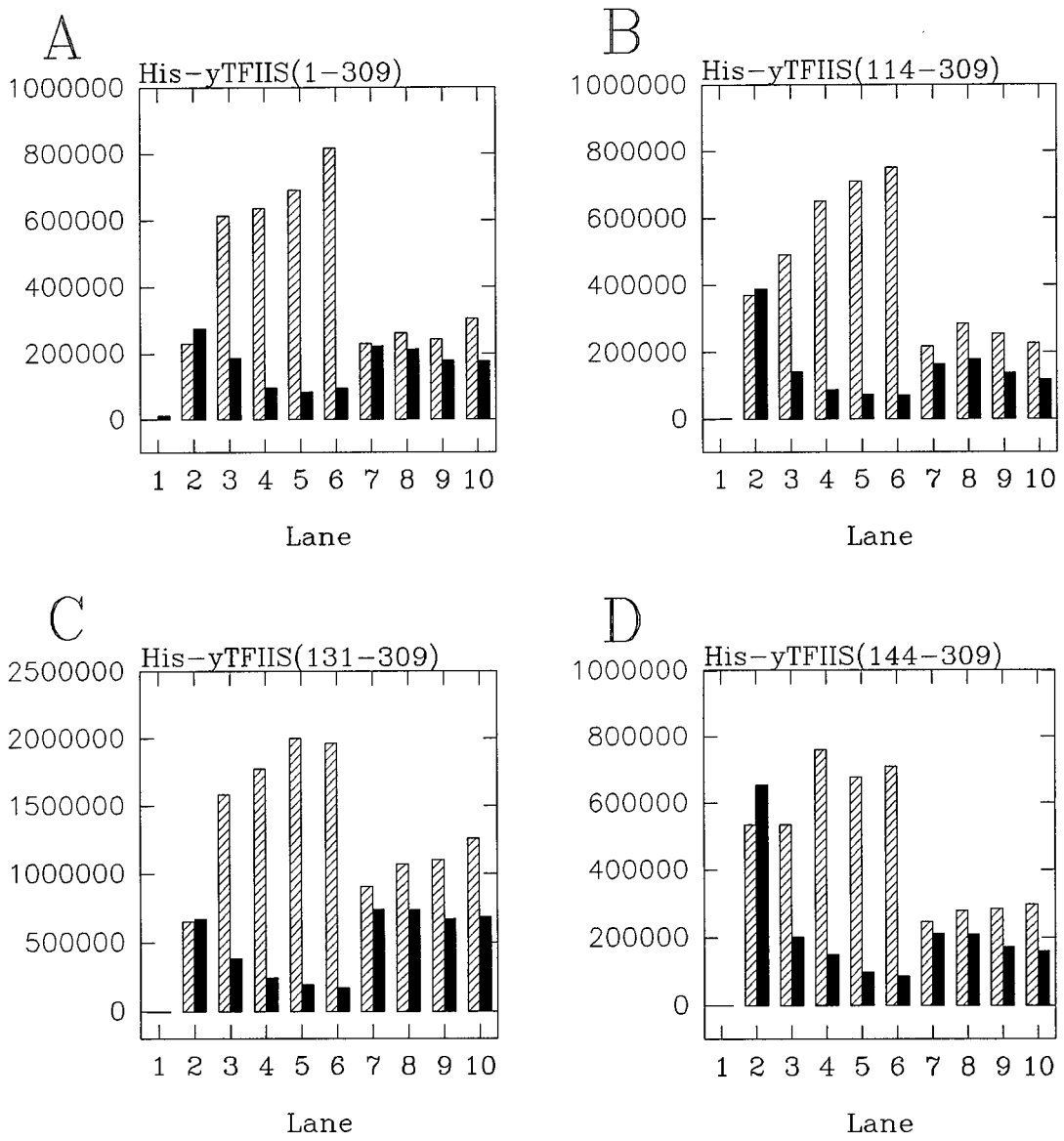


Figure 12 - Time Course of Transcript Cleavage Stimulated by Yeast TFIIS

Isolated ternary complexes were formed with RNA polymerase IIa as described in Materials and Methods. Lane 1: RNAs from isolated complexes without further incubation; Lanes 2-11: Truncated TFIIS was added to isolated ternary complexes and aliquots of the reaction were stopped following incubation at 30°C for 10", 1', 2.5', 5', 7.5', 10', 12.5', 15', 17.5', or 20' after TFIIS addition; Lane 12: Nucleotides (0.8 mM) were added to complexes that had been incubated with TFIIS for 20 sec. Incubation continued at 30°C. The reaction was stopped 20' after nucleotide addition. Lane 13: RNAs from isolated ternary complexes (Lane 1) incubated 20' at 30°C with the addition of nucleotides (0.8 mM). Lane 14: RNAs from isolated ternary complexes (Lane 1) incubated 20' at 30°C with the addition of H_2O instead of nucleotides. The arrows on the left side of the autoradiogram indicate the cleavage products considered in the comparison of the ability of each TFIIS variant to stimulate transcript cleavage.

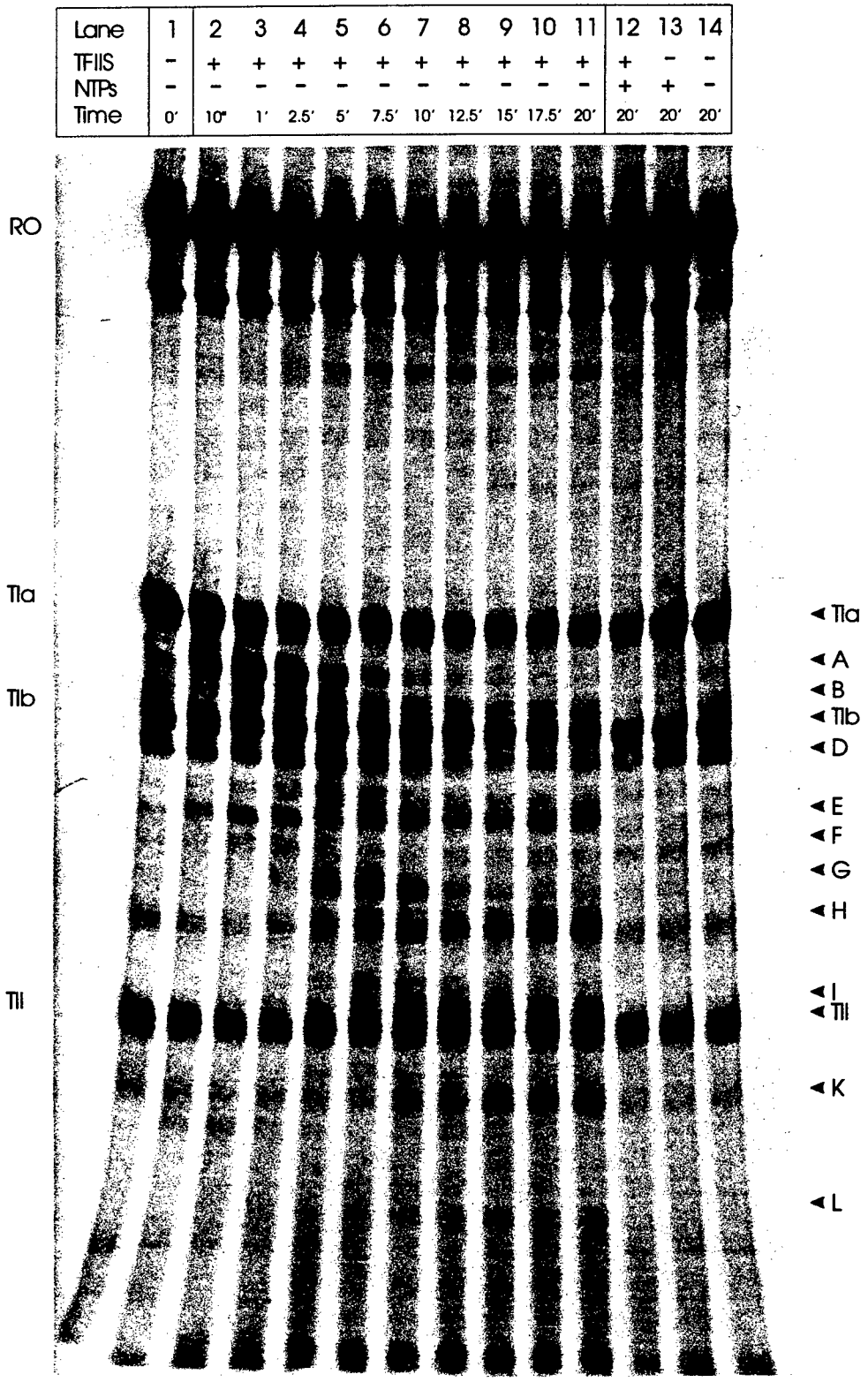

Lane 1 2 3 4 5 6 7 8 9 10 11 12 13 14
TFIIS - + + + + + + + + + + + - -
NTPs - - - - - - - - - - - + + -
Time 0' 10" 1' 2.5' 5' 7.5' 10' 12.5' 15' 17.5' 20' 20' 20' 20'
RO
TIa
TIb
TII
◄ TIa
◄ A
◄ B
◄ TIb
◄ D
◄ E
◄ F
◄ G
◄ H
◄ I
◄ TII
◄ K
◄ L

Figure 13 - Promoter Regions of *RNA15* and *PPR2*

The sequence between the open reading frames of the *RNA15* and *PPR2* genes is shown. Both ORFs are indicated with boxes; an arrow indicates the direction of transcription for each. Putative TATA boxes are indicated by boxes. Putative binding sites for the proteins GCN4, ABF1, and GRF2 are indicated.

Sequence of the Promoter Region Between RNA15 and PPR2

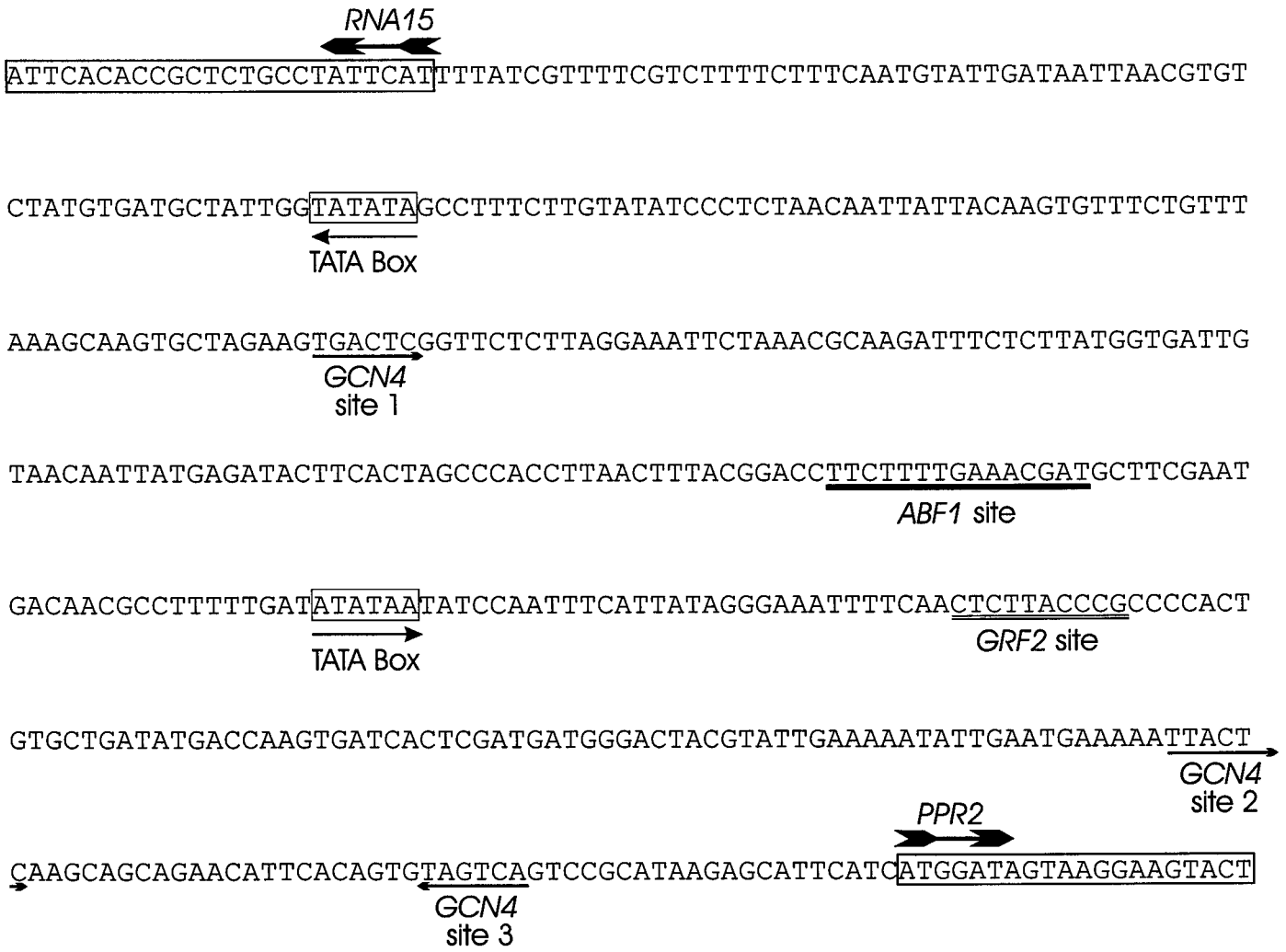


Figure 14 - Alignments of TFIIS Proteins from *S. cerevisiae* and HeLa

A. The sequences of the HeLa and *S. cerevisiae* proteins are shown as aligned by Williams and Kane (1995). That alignment included the predicted amino acid sequences of TFIIS from HeLa, human kidney, mouse Ehrlich ascites tumor, mouse liver 1, mouse liver 2, *D. melanogaster*, *S. cerevisiae*, and *S. pombe*. Only one section of the HeLa and *S. cerevisiae* sequences are shown here. Amino acids that are identical or chemically similar in all eight sequences are highlighted in bold. Identical amino acids that are indicated with a *. Chemically similar amino acids are indicated with a +. The double underlines indicated the cysteines of the zinc-ribbon domain (Qian et al., 1993a). The dotted underline indicates the region of human mutant 1 (Ciprés-Palacín and Kane, 1995). The single underline indicates four amino acids of the yeast protein not aligned with residues of the mammalian sequences in the Williams and Kane alignment. **B.** The minimal sequences known to be active for the HeLa (Ciprés-Palacín and Kane, 1995) and *S. cerevisiae* (Nakanishi et al., 1995) proteins are shown aligned without the gap in the HeLa sequence. The symbols are as described for Panel A and are based on the Williams and Kane alignment.

Alignments of TFIIS Proteins from *S. cerevisiae* and HeLa

A.

```
                                  145 149
S.c. : 127   VSTKPRNSKNDGVDT AIYHHKLRDQVLKALYDVLAKESEHPPQSILHTAKAIESEMNKVNNCD 189
HeLa : 131   PRAP         STSDSVRLKCREMLAAALRTGDDYIAIGA     DEEELGSQIEEAIYQEIR 181
                           +  1   +  **                    +   +
```

B.

```
S.c. : 147                    LRDQVLKALYDVLAKESEHPPQSILHTAKAIESEMNKVNNCD 189
HeLa : 131           PRAPSTSDSVRLKCREMLAAALRTGDDYIAIGADEEELGSQIEEAIYQEIR 181
                             +*  1   +   *                +   +

S.c. : 190  TNEAAYKARYRIIYSNVISKNNPDLKHKIANGDITPEFLATCDAKDLAPAPLKQKIEEIA 249
HeLa : 182  NTDMKYKNRVRSRISNLKDAKNPNLRKNVLCGNIPPDLFARMTAEEMASDELKEMRKNLT 241
                 *+ * *    *+    ** *+       +     +     ++*    +

S.c. : 250  KQNLYNAQGATIERSVTDRFTCGKCKEKKVSYYQLQTRSADEPLTTFCTCEACGNRWKFS 309
HeLa : 242  KEAIREHQMAKTGGTQTDLFTCGKCKKKNCTYTQVQTRSADEPMTTFVVCNECGNRWKFC 301
               +   * *    + **   *+*** +  +* *+*********+***  *  *******
```

Chapter 4

Feasibility Study on The Use of the Two-Hybrid Technique to Identify Proteins Which Interact with Yeast TFIIS *in vivo*

INTRODUCTION

In the previous chapters, I demonstrated that the P37 protein from *S. cerevisiae* is a member of the TFIIS class of transcription elongation factors. I described deletion analysis of the yeast TFIIS protein, both *in vitro* and *in vivo*. This analysis demonstrated that the amino-terminal 143 amino acids of the yeast protein are not required for any known activity of TFIIS: 1) stimulation of read-through of an intrinsic block to elongation by yeast RNA polymerase II; 2) stimulation of cleavage of the nascent RNA transcript by stalled elongation complexes of yeast RNA polymerase II; and 3) complementation of sensitivity to the drug 6-azauracil conferred by deletion of the *PPR2* gene. In addition, this analysis demonstrated that subunits four and seven and the carboxyl-terminal domain of the largest subunit of RNA polymerase II are not required for TFIIS to interact with the polymerase *in vitro* and stimulate either of the first two activities listed above.

The goal of the work reported in this chapter was to elucidate the role of TFIIS *in vivo* by identifying proteins which interact with TFIIS in cells. The identification of these proteins may provide insight into the function or regulation of TFIIS activity. TFIIS activity could be regulated by proteins which bind to or modify the elongation factor to prevent its interaction with the polymerase. Alternatively, a regulatory protein might bind to the polymerase and occlude the binding of TFIIS. Although TFIIS functions as a single peptide *in vitro*, this interaction might be mediated or stabilized by other proteins *in vivo*. To identify proteins which interact with the TFIIS protein in *S. cerevisiae*, I chose to use the two-hybrid technique of S. Fields (Fields and Song, 1989; Chien et al., 1991).

The extensive biochemical characterization of TFIIS from many organisms demonstrates that its function is highly conserved (reviewed in Kassavetis and Geiduschek, 1993; Reines, 1994). TFIIS interacts with RNA polymerase and stimulates cleavage of the nascent transcript to promote read-through of blocks to elongation. Although the polymerase is sufficient to catalyze this reaction under some conditions (Rudd et al., 1994; R. Weilbaecher, unpublished results), TFIIS stimulates the rate of the cleavage reaction. Cleavage of the nascent transcript is proposed to be an essential step in stimulating stalled polymerase complexes to resume elongation, possibly by stimulating a conformational shift in the polymerase which realigns the enzyme's catalytic site with the 3'-end of the growing transcript and allows cleavage (Reines, 1992, 1994; Chamberlin, 1994; Ciprés-Palacín and Kane, 1994, 1995).

The details of how TFIIS may stimulate transcription elongation by RNA polymerase II *in vivo* are currently unknown. Although TFIIS can stimulate *in vitro* read-through of blocks to elongation in genes including the human H3.3 gene (Reines et al., 1989), the c-myc gene (R. Landick, personal communication), the adenosine deaminase gene (Kash et al., 1993), the murine ornithine decarboxylase gene (Shor et al., 1995), and in the adenovirus genome (Wiest et al., 1992), there is no evidence that TFIIS activity regulates the expression of genes *in vivo*. There is evidence that some TFIIS genes have tissue-specific patterns of expression in humans (Weaver, 1995) and in rodents (Xu et al., 1994), but involvement of TFIIS in the control of tissue-specific expression of other genes has not been demonstrated. Since it is clear that TFIIS can interact directly with purified RNA polymerase II in the absence of any of the general initiation factors or components

of the RNA polymerase II holoenzyme, it is difficult to envision a role for TFIIS in the regulation of specific genes.

Deletion of the genes encoding the TFIIS proteins in *S. pombe* or in *S. cerevisiae* (*PPR2*) confers only a mild phenotype (Williams and Kane, 1995; Hubert et al., 1983; Ruet et al., 1990). The cells are viable, but sensitive to the drug 6-azauracil, which causes drops in the cellular pools of UTP and GTP (Exinger and Lacroute, 1992). Additionally, a mild defect in meiotic recombination has been reported (Clark et al., 1991). Amino-terminally truncated TFIIS proteins which are active *in vitro* are able to complement the 6-AU sensitivity of Δppr2 strains, but TFIIS mutants inactive *in vitro* do not complement (Nakanishi et al., 1995). The inseparable correlation of the *in vitro* and *in vivo* activities of TFIIS strongly suggests that the function of this protein in cells is to assist the elongation reaction of RNA polymerase II.

Although it has become obvious in recent years that the initiation reaction can influence the elongation competence of RNA polymerase, these studies do not suggest a role for TFIIS. Clearly other protein factors are involved. Promoters such as that for the mammalian c-myc gene can promote the formation of elongation "incompetent" RNA polymerase II complexes which pause near the promoter and do not transcribe the entire gene (Krumm et al., 1992, 1993, 1995). When expression of the gene is activated, the polymerases which initiate from the c-myc promoter become elongation "competent". A plausible idea is that transactivators acting through TFIIH may stimulate promoter clearance and convert these complexes from elongation "incompetent" to "competent" forms (Goodrich and Tjian, 1994; M. Groudine, personal communication). Studies of the initiation reaction have not yet begun to address how, or even whether, these promoter

effects affect the ability of TFIIS to interact with the polymerase during the elongation phase.

Possibly TFIIS does not regulate specific genes but is required as a general, non-specific elongation factor which prevents polymerase complexes from stalling terminally within a gene. This hypothesis predicts that highly transcriptionally active tissues might require more TFIIS activity. The speculation that TFIIS expression may be stimulated by the general amino acid control transcriptional activator GCN4 in *S. cerevisiae* (Chapter 3 of this work) is consistent with this hypothesis as well.

To identify proteins which interact with TFIIS *in vivo*, I chose to use the yeast two-hybrid technique. It has proven to be a highly successful method with a variety of useful applications. It is a powerful way to characterize protein:protein interactions *in vivo*. It has been used to demonstrate protein:protein interactions between known proteins, notably SIR4-SIR4 homodimers (Chien et al., 1991). Other examples include demonstration of the participation of several of the RAD proteins in formation of a complex required for repair of double-strand breaks (Hays et al; 1995), the interaction of the GCN1 and GCN20 proteins required to activate the GCN2 kinase (Vazquez de Aldana et al., 1995), the formation of heterodimers and homodimers of human nuclear lamin proteins (Ye and Worman; 1995) and the interaction of the cellular prion protein with the Bcl-2 protein (Kurschner and Morgan; 1995). After a protein:protein interaction has been determined, the two-hybrid system provides a powerful method for mapping the interaction to specific domains or residues in the proteins. Domains of the 70 kDa subunit of TFIIIB which mediate, either positively or negatively, the interaction

with the tau 131 subunit of TFIIIC were mapped in this manner (Chaussivert et al., 1995).

One of the most powerful applications of the two-hybrid technique has been the identification of and isolation of genes encoding novel proteins which interact with known proteins or even with DNA sequences (Fields and Sternglanz, 1994). New yeast genes have been cloned in this manner including SNF11, a protein which interacts with the SNF-SWI complex (Treich et al., 1995), and FIP1, a protein which interacts with poly(A) polymerase (Preker et al., 1995). The two-hybrid system indicated that a putative calcium binding protein, identical to ERC-55, may interact with the oncogenic E6 protein from human papillomaviruses (Chen et al., 1995). The technique has proved immensely valuable in investigating proteins from other eukaryotic systems. A variety of novel genes have been isolated. Examples include a new mouse member of the E2F family, a group of proteins implicated in the regulation of cellular proliferation through interaction with the retinoblastoma protein (Buck et al., 1995), a human protein EB1 which interacts with the APC protein, mutations in which are associated with both sporadic and familial colorectal cancer (Su, et al; 1995), a gene for a novel human protein which interacts with the tumor necrosis factor receptor (Boldin, et al; 1995), and a receptor protein for bone morphogenic factors (Liu, et al; 1995). This sampling is but a fraction of the successful applications for this technique.

In this work, I investigated the feasibility of using the two-hybrid technique to identify proteins which interact with TFIIS *in vivo* in *S. cerevisiae*. Preliminary screens were conducted to test the selection technique. The work described here indicates that the two-hybrid system may provide a useful tool for the elucidation of proteins which interact with the transcription elongation factor TFIIS *in vivo* in

S. cerevisiae. A variety of controls demonstrate that, despite the direct interaction of TFIIS with RNA polymerase II, a fusion protein with an active TFIIS protein joined to the DNA-binding domain of the GAL4 protein does not activate transcription of two GAL4-dependent reporter constructs. The identification and characterization of proteins which interact with this transcription elongation factor may significantly advance our understanding of the role of this factor *in vivo.*

MATERIALS AND METHODS

Construction of pKC14a-1

In order to subclone the *PPR2* open reading frame (ORF) in frame with the DNA binding domain of GAL4, it was necessary to design primers for the polymerase chain reaction (PCR). The primer P37-3D (5'-CGGGATCCCGATGGA-TAGTAAGGAAGTACTG-3') encodes a BamHI site upstream of the first ATG codon of the *PPR2* ORF, such that the ORF is subcloned in frame with the GAL4 DNA binding domain in the plasmid pAS1. The primer p37-4U (5'-GGCGCGTCGACAT-GCATAGAAGATTAATAAAG-3') adds a SalI site approximately 300 base pairs downstream of the *PPR2* ORF to facilitate directional cloning. Using these primers, the region containing the *PPR2* ORF was amplified from pKC3 (see Chapter 3) using Taq DNA polymerase in standard PCR conditions [20 mM Tris-Cl, pH 8.5; 3.5 mM $MgCl_2$; 50 mM KCl; 0.1% Triton X-100; 0.8 mM deoxynucleotides (dNTPs); 1 µM each primer] in a 100 µl reaction for 30 cycles [Denaturation at 94°C for 1 minute, annealing at 47°C for 1 minute, and extension at 72°C for 3 minutes]. The product of the PCR was purified from remaining primers and dNTPs using a Millipore spin filter. The filter retains the product DNA due to its size, while primers and nucleotides pass through. The product was washed 3 times with TE, spinning 10 minutes at 4500 rpm (2000 g) in the angle rotor of a BioFuge B, and then eluted from the filter in a total volume of 30 µl TE.

The product of the PCR and pAS1 plasmid DNA were digested with SalI (NEB) and then BamHI (NEB), gel purified using a 1% agarose/TAE gel, and extracted from the agarose using Gene Clean. The two DNAs were ligated together using T4 DNA ligase (USB) using standard ligation conditions [40 mM Tris-Cl, pH 7.5; 10 mM $MgCl_2$; 10 mM DTT; 50 µg/ml acetylated BSA; 0.67 mM ATP; 0.01 units

T4 DNA ligase (1 µl from a dilution of 1 unit of USB T4 DNA ligase diluted 100-fold in ligase diluent - 20 mM Tris-Cl, pH 7.5; 1 mM EDTA; 5 mM DTT; 60 mM KCl; 50% glycerol). The ligation mixtures were electroporated into XL-1Blue cells which were plated on LB-ampicillin plates. DNA was prepared from transformants and restriction analysis was used to determine which clones contained the desired insert fragment. One of these transformants was selected and the PCR-generated insert fragment was sequenced. One base change was found in the third position of an alanine codon. Since this mutation is silent, the plasmid was named pKC14a-1.

Preparation of *E. coli* cells for Electroporation

A 4 ml culture of LB was inoculated with *E. coli* and shaken at 37°C overnight. A 300 ml LB culture was inoculated with 3 mls of the overnight culture and shaken at 37°C until the OD_{600} reached approximately 0.5. The flask containing the cells was chilled on ice for 15-30 minutes. The cells were harvested by spinning for 15 minutes at 5 K rpm in a Sorvall GSA rotor. The cells were washed in 200 mls cold 1 mM HEPES, pH 7 and pelleted again as before. The cells were washed a second time in 150 mls cold 1 mM HEPES and pelleted again. The cells were washed one time with 6 mls cold 10% glycerol and pelleted by spinning for 15 minutes at 7 K rpm in a Sorvall SS-34 rotor. The entire cell pellet was brought to a final volume of 1 ml with cold 10% glycerol. Aliquots of 40 µl cells were placed in Eppendorf tubes on ice and mixed with DNA in a volume of 1-2 µl. The cell-DNA mix was transferred to chilled 0.2 cm electroporation cuvettes and electroporated using a Gene Pulser [set at 25 µF, 2.5 kV, and 200 Ω]. Immediately after pulsing, 1 ml cold LB or SOC [LB + 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM

$MgSO_4$, 20 mM glucose] was added. The cells were transferred to small glass tubes, grown for 1 hour at 37°C, and plated on the appropriate selective media.

TDMN Protocol for Sequencing

Sequencing was performed using the TDMN procedure for the Sanger sequencing method (Del Sal et al., 1989), as modified by D. Mytelka (Mytelka, 1995). In brief, the TDMN protocol is as follows. The supercoiled plasmid to be sequenced is mixed with a primer (0.4 µM) and 0.1 M NaOH in a total volume of 10 µl and incubated 10 minutes at 68°C. Immediately upon removal from the 68°C incubation, 4 µl TDMN [0.28 M TES, 0.2 M NaCl, 0.08 M $MgCl_2$, 0.12 M HCl, 50 mM DTT; made fresh from 10X stock solutions] solution is added and the primer anneals to the plasmid while the mixture is incubated at room temperature (RT) for at least 10 minutes, but not to exceed 2 hours. During this time, 2.5 µl USB termination mixes and 3.5 µl of 5.5 M betaine (Mytelka, 1995) are aliquotted to microtiter wells and pre-warmed to 37°C. After annealing, 5 µl of Labeling Cocktail [USB dGTP sequencing labeling mix; [α^{35}S]-dATP; Sequenase, mixed per USB's instructions for Sequenase] is added and this labelling reaction is incubated at RT for 5 minutes. After the 5 minute RT incubation, 3.5 µl of the labeling reaction is aliquotted to each of four wells containing one of the four different termination mixes and the reaction is incubated 5 minutes at 37°C. At this time, 4 µl formamide stop mix [90 mM Tris-borate, pH 8.0, 2 mM EDTA, 80% formamide, 0.13% bromophenol blue, 0.13% xylene cyanol, 0.06% SDS] is added, the reactions are heated 2.5 minutes at 95°C to denature the DNA and chilled on ice. An aliquot of each reaction was resolved on a 40 cm 6% acrylamide/8.3 M urea gel run at 35W. The gel was rinsed with 10% acetic acid, 12% methanol, dried and exposed to film.

Transformation of Supercoiled Plasmids into *Saccharomyces cerevisiae*

S. cerevisiae cells were made competent for transformation as described in Chapter 3, except after resuspension, the cells were spread on the appropriate selective media for the plasmid. Generally, if two plasmids were desired in a strain, the transformations were done sequentially rather than at the same time.

Growth Media for Yeast Strains

Liquid media was prepared in 0.5 L aliquots as follows: 3.35g Difco yeast nitrogen base w/o amino acids, 0.56g "almost complete" amino acid mix (composition described below), the appropriate supplemental mix, 10 g dextrose, and H_2O to a final volume of 500 mls were mixed. The supplemental mixes were as follows: for SC - tryptophan, 0.325 g uracil drop-out mix and 0.010 g uracil; for SC - leucine, 0.205 g leucine drop-out mix and 0.015 g tryptophan; for SC - histidine, leucine, tryptophan, 0.175 g met, lys, ura mix; and for SC - histidine, leucine, tryptophan, uracil, 0.165 g met, lys mix, amounts given per 0.5 L. Liquid media was sterilized by filtration.

"Almost complete" amino acid mix is composed of 40 mg/L adenine-HCl, 20 mg/L arginine, 30 mg/L tyrosine, 30 mg/L isoleucine, 50 mg/L phenylalanine, 100 mg/L glutamic acid, 100 mg/L aspartic acid, 200 mg/L threonine, 400 mg/L serine, and 150 mg/L valine. Uracil drop-out mix is composed of 150 mg/L methionine, 180 mg/L lysine, 260 mg/L leucine, and 60 mg/L histidine. Leucine drop-out mix is composed of 150 mg/L methionine, 180 mg/L lysine, 20 mg/L uracil, and 60 mg/L histidine. Met, lys, ura mix is composed of 150 mg/L methionine, 180 mg/L lysine, 20 mg/L uracil and Met, lys mix is identical except for the absence of uracil.

Solid media lacking or containing 50 mM 3-aminotriazole (3-AT) was prepared as follows. A solution of 3 M 3-aminotriazole (Aldrich) was prepared in H_2O and filter sterilized. A solution of 40% dextrose was prepared and sterilized by autoclaving. SC - histidine, leucine, tryptophan or SC - histidine, leucine, tryptophan, uracil was prepared as described above except 20 g/L agar (Bacto) was added, the dextrose was omitted, and the mixture was autoclaved and cooled to about 50°C. To the cooled agar mixture was added, 40% sterile glucose to a concentration of 2%, and the appropriate volume of 3 M 3-aminotriazole to give the desired final concentration (see text and Figure Legends). The mixture was stirred and poured into Petri dishes to solidify. Alternatively, 3-aminotriazole was spread on solid minimal media lacking histidine to give the final concentration desired.

Filter Assay for β-Galactosidase Activity in Yeast Cells

This protocol accompanied literature from the Elledge laboratory on the use of the Y153 strain and is originally from P. Chevray. The cells to be assayed were streaked on an appropriate selective media and grown at 30°C until colonies were big enough to lift, usually three days. For each plate to be assayed, a 7 cm circular Whatman filter was soaked in 2 mls Z-Buffer + 40 mM βME (Kodak) and 0.335 mg/ml X-Gal, made fresh from Z-Buffer [60 mM Na_2HPO_4 (dibasic), 40 mM NaH_2PO_4 (monobasic), 10 mM KCl, 1 mM $MgSO_4 7H_2O$; pH 7.0], 14 M βME, and 40 mg/ml X-Gal in N,N-dimethylformamide. A sterile 7 cm circular Whatman filter was placed on a plate until completely wetted. It was then lifted off and placed in liquid nitrogen until completely frozen. The filter was removed from the nitrogen, allowed to thaw at RT, and placed cell side up on a filter saturated with Z-Buffer +

βME and X-Gal and left overnight to develop. Generally blue color began to appear within 1.5 - 2 hours.

Library Transformation Protocol

This protocol for high efficiency transformation (~30,000 colonies/μgram), based on that of Schiestl and Gietz (1989), was recommended by J. Gowen-Cooke. A 10 ml SC - trp culture was inoculated with Y166 + pKC14a-1 and shaken at 30°C overnight. A 100 ml SC - trp culture was inoculated with an aliquot of this overnight culture to an OD_{600} of approximately 0.1 and shaken at 30°C until the OD_{600} reached between 0.8 to 0.9. The cells were spun down 5 minutes at 5 K rpm in a Sorvall GSA or (GS3) rotor. They were washed in a volume of LiOAc/TE (100 mM LiOAc, 10 mM Tris-Cl, pH 7.5; 1 mM EDTA) equal to the original culture volume and collected by spinning as before. The pellet was vortexed to resuspend the cells in the residual liquid and 100 μl aliquots were placed in Eppendorf tubes. To the cells were added, 10 μl 10 mg/ml carrier DNA, between 0.5 and 5 μg library DNA (YL1, YL2, or YL3; Chien et al., 1991) or control plasmid DNA (usually pNI12; Fields and Song, 1989), and 600 μl PEG/LiOAc/TE (40 % PEG, MW 3350; 100 mM LiOAc, 10 mM Tris-Cl, pH 7.5; 1 mM EDTA). The cells were incubated 2-4 hours in this mixture at 30°C, either with periodic inversion or with constant rotating. They were then heat shocked for 15 minutes at 42°C, transferred to 15 ml Corning orange cap polypropylene tubes and washed two times with 10 mls sterile H_2O. The cells were resuspended in 6 mls SC - his, leu, trp media. Small aliquots of 10 μl, 20 μl, or 200 μl were plated on SC - leu or on SC - his, leu, trp to determine transformation efficiency. The remainder of the cells were stored at 4°C

until transformation efficiency was calculated. The remaining cells were plated on SC - his, leu, trp, ura to achieve 1000-2000 colonies per 100 millimeter plate.

Preparation of Genomic DNA from *Saccharomyces cerevisiae*

Yeast genomic DNA was prepared by the method of Winston. Briefly, 10 ml cultures of YPD were inoculated with yeast cells and shaken at 30°C at least 18 hours. The cells were collected by centrifugation 5 minutes at 3 K rpm in an SS-34 Sorvall rotor. The cells were resuspended in 0.5 mls water, transferred to an Eppendorf tube, and pelleted with a brief spin. The supernatant was discarded and the cells were resuspended in the residual liquid. To the cells was added, 200 µl of Solution A (2% Triton X100, 1% SDS, 100 mM NaCl, 10 mM Tris-Cl, pH 8.0, 1 mM EDTA), 100 µl neutral phenol, 100 µl SEVAG, and 0.3 g acid washed glass beads. After vortexing this mixture vigorously for 3 minutes, 200 µl of TE, pH 8.0 was added and the tube was spun in the microfuge for 5 minutes. The aqueous supernatant above the organic layer, also containing cell debris and glass beads, was transferred to a new tube. The DNA was precipitated by the addition of 900 µl cold ethanol (mixed by inversion) and pelleted by spinning 2 minutes in the microfuge. The DNA was dissolved in 400 µl H_2O, 30 µg DNase-free RNaseA was added, and the tube was incubated 5 minutes at 37°C. The DNA was ethanol precipitated again, dried and dissolved in 50 µl TE, pH 8.0. A 10 µl aliquot of this preparation is generally 2-4 µgs DNA, an appropriate amount for one lane of a Southern blot.

Lithium Chloride/Triton X-100 Mini-DNA Preparations from *S. cerevisiae*

DNA from *S. cerevisiae* was prepared using a lithium chloride/Triton X-100 mini-preparation (Caldwell and Becker, 1994; Ward, 1990). 2 ml SC - leu cultures were inoculated with putatively positive isolates from the preliminary YL1 screen and shaken overnight (at least 16 hours) at 30°C. Cells were harvested by centrifugation in Eppendorf tubes and the supernatant was discarded. To each tube was added 100 µl LiCl/Triton solution [2.5 M LiCl; 50 mM Tris-Cl, pH 8.0; 4% Triton X-100; 62.5 mM EDTA], 50 µl neutralized phenol, 50 µl SEVAG (24 : 1 :: chloroform : isoamyl alcohol), and 0.2 g acid-washed glass beads. These mixtures were vortexed vigorously for 2 minutes and spun 2 minutes in a microfuge. The cell debris formed a compact pellicle above the organic layer. The supernatant was transferred to a new tube containing 250 µl cold ethanol and the DNA was precipitated by spinning at least 15 minutes at 4°C in a microfuge. The pellet was washed with 70% ethanol, dried and dissolved in 20 µl H_2O.

Preparation of Competent MH1 *E. coli* cells by Rubidium Chloride Method

The MH1 strain (Hall et al., 1984) and this transformation method were recommended by K. Arndt (Univ. of Pittsburgh) since it is particularly effective for plasmid rescue when using DNA prepared from yeast by the method of Winston. A 20 ml LB culture was inoculated with MH1 *E. coli* cells and shaken at 37°C overnight. A 5 ml aliquot of this culture was used to inoculate a 500 ml culture of LB + 20 mM $MgCl_2$, which was shaken at 37°C until the OD_{550} reached, but did not exceed, an optical density of 0.4. The cells were chilled on ice, transferred to 2 chilled, sterile GS3 bottles, and collected by spinning 5 minutes at 5 K rpm in a Sorvall GS3 rotor. The supernatant was discarded and the cell pellet in each tube

was kept on ice and gently resuspended in 75 mls cold TfB I [30 mM KOAc; 50 mM $MnCl_2$; 100 mM $RbCl_2$; 10 mM $CaCl_2$; 15% glycerol; brought to pH 5.8 with 0.2 N HOAc and sterilized by filtration]. The cells were allowed to sit on ice an additional 2.5 minutes after they were completely resuspended and were then pelleted by spinning 5 minutes at 5 K rpm in the GS3 rotor. The supernatant was discarded and each cell pellet was resuspended in 10 mls cold TfB II [10 mM MOPS, pH 7.0; 75 mM $CaCl_2$; 10 mM $RbCl_2$; 15% glycerol] by gentle pipetting. The cells were aliquotted to Eppendorf tubes (approximately 100 µl per tube) and frozen in dry ice/ethanol. This freeze step is absolutely essential for induction of competence.

For transformation, the appropriate number of aliquots was thawed on ice for 10 minutes to 1 hour. Transforming DNA was added in less than 25 µl (generally 1 µl or less for DNA prepared by the Winston method) and the cell-DNA mix was incubated on ice for 20 minutes. The cells were then heat shocked for 2 minutes at 37°C and 1 ml of LB was added to the Eppendorf tube. The tubes were capped tightly, placed in a flask, and shaken at 37°C for 1 hour. They were then plated on LB-ampicillin plates. This process routinely generated cells with an efficiency of 3×10^8 colonies/µg and cells stored at -80°C could be used at least six months later with negligible drop in efficiency. Generally, very little DNA was needed, and in many cases less was better. Often 0.1 µl worked better than 1 µl or more of the DNA prep.

Retransformation into DH10B *E. coli* Cells for DNA Preparation

In order to prepare better quality DNA for sequencing and restriction analysis, mini prep DNA of library plasmids prepared from MH1 cells using the

lithium chloride/Triton X-100 mini-preparation for DNA from *E. coli* described in Chapter 3 was transformed into DH10B *E. coli* cells, prepared and transformed as described in Chapter 3. pYL2 library plasmids were prepared by the alkaline lysis method described in Chapter 3, except a 20 ml culture was grown and all subsequent volumes were scaled accordingly. DNA prepared in this manner from DH10B cells was suitable for restriction or sequence analysis. DNA prepared in this manner from MH1 cells was suitable for use as transforming DNA but produced variable results in restriction analysis and was unsuitable for sequencing.

Sequencing and Sequence Analysis of pYL2 Plasmids

The pYL2 plasmids were sequenced using a primer GAL4-ENDD (5'-GATG-AAGATACCCCACCAAACCC-3') with either of two methods. Two plasmids, pYL2-1B and pYL2-2, were sequenced by the Barker Hall DNA Sequencing Facility as described in Chapter 3. The remaining six plasmids were sequenced using the Mytelka modification of Sanger sequencing described above. The nucleic acid sequences were used as query sequences to search the non-redundant nucleic acid databases with the blastn program (Altschul et al., 1990). The sequences were translated in all three possible frames and each peptide was used as a query sequence to search the non-redundant protein databases with the blastp program (Altschul et al., 1990).

RESULTS

Purpose

The goal of this work was to determine if the two-hybrid technique could be utilized to identify proteins which interact with TFIIS *in vivo* in *S. cerevisiae.* The chosen two-hybrid fusion scheme utilizes the independent and separable activities of the DNA-binding domain (DBD) and the transcriptional activation domain (AD) of the GAL4 protein. The amino-terminal portion of GAL4 is sufficient to bind DNA, but does not activate transcription. The carboxyl-terminus of GAL4 does not bind DNA, but is sufficient for transcriptional activation when localized to a promoter. When not covalently attached, the two domains do not interact (Ma and Ptashne, 1987). However, if the activation domain is brought to the DNA by some means other than covalent attachment to the GAL4-DBD, it is sufficient to activate transcription. Thus the basic idea, diagrammed in Figure 1, is to construct a fusion protein composed of the GAL4-DBD and the protein of interest, the "bait", TFIIS in this case. A plasmid expressing this fusion protein is transformed into strains containing one or more GAL4-dependent reporter genes. The "bait" protein is localized to the promoter of the reporter by the GAL4-DBD, but does not activate its transcription. The introduction of a second fusion protein, composed of the GAL4-AD and a protein which interacts with the "bait", will bring the GAL4-AD to the promoter. Thus the activity of the GAL4 transcriptional activator is dependent on a protein:protein interaction between the two proteins fused to the DBD and AD portions of GAL4. The use of an expression library fused to the GAL4-AD allows the identification of proteins which interact with the protein of interest, here TFIIS.

Construction of Plasmid Encoding Fusion of *GAL4* DNA Binding Domain to TFIIS

A variety of plasmids are available for the construction of fusions to the DNA binding domain of GAL4. They differ primarily in the restriction enzyme sites available for cloning the open reading frame (ORF) of interest in frame with that of the GAL4-DBD and in the selectable marker. The plasmid pAS1 was chosen because it encodes an hemagglutinin (HA) epitope in between the GAL4 DNA-binding domain and the BamHI site used for insertion of the protein of choice. The presence of this epitope would allow the use of antibodies against hemagglutinin to detect expression of the fusion protein. The entire TFIIS ORF was subcloned into this vector (see Materials and Methods for details) and the resulting plasmid named pKC14a-1. Figure 2 shows a diagram of this plasmid and the resulting polylinker sequences after cloning the *PPR2* ORF into pAS1. pKC14a-1 was transformed into two reporter strains, Y153 and Y166.

Detection of Expression of the Fusion Protein from pKC14a-1

Before proceeding, it was essential to confirm that pKC14a-1 was producing the desired fusion protein, which will be referred to as DBD-TFIIS. Although no antibodies against yeast TFIIS were available at the time, antibodies against either the hemagglutinin (HA) epitope or against the DNA binding domain of the GAL4 protein (amino acids 1-147) were available. Either of these antibodies should detect the GAL4 fragment with the HA epitope expressed by pAS1 or the fusion encoded by pKC14a-1.

Expression of a fusion protein from pKC14a-1 was analyzed in the Y153 strain. Protein extracts from Y153 containing either pAS1 or pKC14a-1 were prepared and resolved on a 12% SDS-polyacrylamide gel. The proteins were

transferred to a nitrocellulose filter and the blot was probed with a polyclonal mouse anti-HA primary antibody. The secondary antibody was a sheep anti-mouse IgG conjugated to horseradish peroxidase, permitting detection of proteins by chemiluminescence (all done essentially as described in Methods for Chapter 3). Western analysis suggested that cells bearing the pKC14a-1 plasmid were producing an intact fusion of the GAL4 DNA binding domain and the TFIIS protein.

As technical problems made the Western analysis somewhat ambiguous, the expression of an intact fusion protein was assayed by a second method. If the pKC14a-1 plasmid expressed TFIIS activity, then it would confer 6-azauracil (6-AU) resistance to cells deleted for the *PPR2* gene. To test complementation, the plasmids pAS1 and pKC14a-1 were transformed into KJY1, a strain deleted for the *PPR2* gene (described in Chapter 3). Transformants were incubated at 30°C on solid minimal media lacking uracil and tryptophan and containing various concentrations of 6-AU. The presence of pKC14a-1 allowed KJY1 to grow in the presence of 6-AU while the vector pAS1 did not (data not shown). The ability of pKC14a-1 to complement the defect conferred by deletion of *PPR2* demonstrates the expression of active TFIIS, very likely as part of a fusion protein since sequencing had indicated the presence of the expected cloning junction.

Reporter Strains

Similarly to the variety of plasmids which have been constructed to make fusions to the DNA-binding domain of *GAL4*, many reporter strains are available. Two strains, Y153 and Y166 (Table 1) were compared to determine suitability for use with pKC14a-1. The Y153 strain contains two integrated reporter constructs,

one expressing lacZ and one expressing *HIS3*. The Y166 strain contains three integrated reporter constructs, expressing lacZ, *HIS3*, or *URA3*. The laboratory of S. Elledge has determined that the *URA3* of Y166 reporter requires two-fold higher levels of activation than the other reporter constructs (S. Elledge, personal communication). All of the reporter genes utilize the *GAL1* promoter or a GAL4-dependent upstream activating sequence (UAS) and are thus dependent on either GAL4 or a two-hybrid interaction reconstituting GAL4 for transcription. The lacZ reporter provides a color screen to detect activation. The *HIS3* and *URA3* reporters provide selections for prototrophy.

There are several advantages to using a strain containing two integrated reporter constructs. In two-hybrid analyses, some false positives are obtained because the activation domain of *GAL4* is fused to a library protein capable of binding to a sequence in the promoter (Bartel et al., 1993). These types of false positives characteristically activate one reporter construct while a positive that requires interaction with TFIIS will activate both reporter constructs (Bartel et al., 1993; Fields and Sternglanz, 1994. The use of a strain containing more than one reporter construct with slightly different promoters facilitates detection of this type of artifact.

Assays for Activation of the Three Integrated Reporter Constructs

Before proceeding to a library screen, it was necessary to determine that the fusion of TFIIS to the DNA binding domain of GAL4 did not activate transcription of GAL4-dependent reporter constructs. If a fusion protein containing the "bait" of interest, in this case DBD-TFIIS, activates transcription of the reporter genes, then it cannot be used in the two-hybrid system.

Activation of the lacZ reporter construct is simple to assay. A sterile Whatman filter is laid onto a plate containing colonies of the strain to be assayed. After allowing the filter to become completely wet, the filter is lifted from the plate, frozen in liquid nitrogen for five seconds, and allowed to thaw at room temperature for several seconds. The filter is then placed, cell side up, on a second filter saturated with the assay buffer containing 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal). Positive colonies turn blue after two or more hours incubation at room temperature.

The assay to test activation of the *HIS3* reporter gene is also straightforward. Since the reporter strains are deleted for the chromosomal *HIS3* locus, the integrated reporter construct is the only copy of the gene in these cells. However, it has been observed that the *HIS3* reporter is expressed at a low constitutive level sufficient for Y153 or Y166 to grow without histidine in the media. The addition of 3-aminotriazole (3-AT), a competitive inhibitor of the *HIS3* gene product (Klopotowski and Wiater, 1965; Struhl and Davis, 1977; Hill et al., 1986), renders this low level of *HIS3* expression insufficient for viability. Thus, the strain cannot grow on media lacking histidine unless the GAL4-dependent *HIS3* reporter is activated to produce the His+ phenotype. Thus the ability of a "bait" fusion protein to activate this reporter is assayed by plating the test strains on media lacking histidine and containing various concentrations of 3-AT. For the Y153 strain, concentrations between 20-50 mM 3-AT prevent growth of cells containing "bait" proteins which do not activate transcription of the reporter (S. Elledge, personal communication).

Testing activation of the *URA3* reporter is the simplest of the three assays. The GAL4-dependent *URA3* construct is the only copy of this gene in the Y166

strain. The test strains are plated on media lacking uracil and incubated at 30°C. If the "bait" protein does not activate transcription to produce the Ura+ phenotype, the cells do not grow.

Comparison of Reporter Activation by DBD-TFIIS in Y153 and Y166

In order to determine which of the two reporter strains to use, the ability of pKC14a-1 to activate each of the reporter genes in Y153 and in Y166 was tested. Although pAS1, the parent plasmid, might be expected to function as a negative control, it has been observed that it activates transcription of both of the reporters in Y153, most likely due to the translation of polylinker sequences as a fusion with the GAL4 DNA-binding domain (S. Elledge, personal communication; data not shown). For this reason, cells containing a plasmid derived from pAS1, called pAS-SNF1, were used as a negative control instead. This plasmid encodes a fusion of the GAL4 DBD to the SNF1 protein (DBD-SNF1); it has been demonstrated that this fusion protein does not activate transcription of GAL4-dependent reporters (Fields and Song, 1989). Two positive controls were used. The plasmid pCL1 contains the entire GAL4 protein expressed from the ADH promoter (Fields and Song, 1989). A fusion of the GAL4 activation domain (AD) to the SNF4 protein (AD-SNF4) is encoded by the plasmid pNI12 and in combination with DBD-SNF1 reconstitutes GAL4 activity through the SNF1-SNF4 interaction (Fields and Song, 1989). The pNI12 plasmid was also transformed into cells containing pKC14a-1 to test the ability of DBD-TFIIS to activate reporters in the presence of an AD-fusion protein on media appropriate for a library screen. It was presumed that TFIIS and SNF4 do not interact.

To test activation of the *HIS3* and *URA3* reporter genes by DBD-TFIIS, the growth of the two strains bearing either pAS-SNF1 or pKC14a-1, with or without pNI12, was compared on a variety of media. Activation of the *HIS3* reporter by the DBD-TFIIS fusion protein was determined by growth on solid minimal media lacking histidine and tryptophan (SC - his, trp) and containing no, 25 mM, 50 mM, or 75 mM 3-aminotriazole. The same concentrations of 3-AT were also used in solid media lacking leucine as well as histidine and tryptophan (SC - his, leu, trp); the absence of leucine selects for maintenance of pNI12. Activation of the *URA3* reporter in Y166 by the DBD-TFIIS fusion was determined by growth on solid minimal media lacking uracil. (Y153 grows on this media since the strain is prototrophic for uracil.) The results of three days growth on SC - his, trp containing 3-AT and on SC - ura are shown in Figure 3.

The *HIS3* reporter looks usable in both reporter strains. The DBD-TFIIS fusion protein activates a slightly higher level of expression of the *HIS3* reporter in both strains than does the DBD-SNF1 protein. However 50 mM 3-AT is sufficient to inhibit growth of Y153 expressing the DBD-TFIIS protein. The Y166 strain appears to be more resistant to 3-AT since, when expressing both DBD-SNF1 and AD-SNF4, it grew better in the presence of 75 mM 3-AT than Y153 expressing both proteins. This suggests that higher 3-AT concentrations might be more effective for this strain. The need for higher concentrations of 3-AT to inhibit basal activity of the *HIS3* reporter in Y166 was even more pronounced after five days of incubation (data not shown). For the cells bearing the pNI12 plasmid, similar results were seen on SC - his, leu, trp media (data not shown).

Both Y166 positive control strains containing GAL4 activity, either reconstituted through the SNF1-SNF4 interaction or expressed from pCL1, grew

robustly in the absence of uracil. Y166 cells expressing either DBD-TFIIS or DBD-SNF1 alone did not grow. At five days, growth of Y166 + pKC14a-1 was barely detectable but it was essentially equivalent to the growth of Y166 + pAS-SNF1 (data not shown). These results indicate that the *URA3* reporter will allow the use of the DBD-TFIIS fusion protein as a "bait" in a screen of an activation domain library to identify proteins which interact with TFIIS *in vivo*.

The ability of the DBD-TFIIS protein to activate the lacZ reporter in each of the two strains was also tested. Y153 and Y166 cells bearing either pKC14a-1 or pAS-SNF1 were grown on solid minimal media lacking tryptophan (SC - trp) and the strains also bearing pNI12 were grown on SC - his, leu, trp. After several days incubation at 30°C, the cells were assayed for β-galactosidase activity. The results are shown in Figure 4. It is obvious that the level of activation of the lacZ reporter in Y153 by the DBD-TFIIS protein is too high to use this reporter even as a secondary screen. Although Y166 bearing only pKC14a-1 developed a mild blue color, when this strain also contained pNI12 blue color was almost undetectable. This confirmed that TFIIS and SNF4 do not interact. Thus in Y166, the lacZ reporter generates a low but acceptable background level of blue color, and can be useful as a secondary screen.

The results of testing the integrated reporter constructs in Y153 and Y166 suggested that the Y166 strain could be used to screen an activation domain library for proteins which interact with TFIIS *in vivo*. The presence of the "bait" protein, DBD-TFIIS, did not activate transcription of the *URA3* reporter construct. Thus, it appeared to be ideal for use as the primary selection for library positives. Either the *HIS3* or lacZ reporters appeared to be usable as a secondary criterion.

The Effects of the Growth Media on Activation of Reporters in Y166

During several experiments assaying for β–galactosidase activity, I had noticed that the results obtained varied somewhat depending on the selective media used. Both the rate of growth and the development of blue color appeared to be affected by the absence or presence of histidine, leucine, tryptophan, and uracil in various combinations in the media. For this reason, I investigated the effects of the composition of the media on the outcome of assays for β–galactosidase activity.

For this set of assays, Y166 cells were grown on solid minimal media lacking histidine, leucine, tryptophan, or uracil, singly or in combinations. The strain Y166 + pKC14a-1, pNI12 will be referred to as the experimental strain and Y166 + pAS-SNF1, pNI12 will be referred to as the positive control strain. The presence of the pNI12 plasmid allowed the cells to grow on all of the selective media to be tested and previous work had indicated no interaction between yeast TFIIS and the SNF4 protein encoded by pNI12. The experimental and positive control strains were streaked on each type of media and incubated at 30°C for 5 days. The results are displayed in Figure 5; the positive control strain is on the top half of each plate and the experimental strain is on the lower half. Subsequently a Whatman filter lift of each plate was made to assay for β–galactosidase activity; these results are shown in Figure 6.

Comparison of the growth of the experimental strain expressing DBD-TFIIS (and AD-SNF4) on these various media corroborates the previous experiment indicating that the TFIIS "bait" protein does not activate the *URA3* reporter. On either SC - ura or SC - his, leu, trp, ura the experimental strain shows little or no detectable growth. However, this strain grew unexpectedly well on SC - leu, trp,

ura. It is not clear why this media would allow growth of Y166 expressing DBD-TFIIS when media lacking only uracil or media lacking all four tested nutrients does not. I had been leaving histidine out of the media to allow testing of the *HIS3* reporter by spreading 3-aminotriazole on the plates. This result provides an additional reason for its omission.

The results of these assays for expression of the lacZ reporter in Y166 indicate that activation by DBD-TFIIS is minimal. When the cells are grown on media lacking uracil as appropriate for a library screen, there is no blue color. Although the selection for the Ura+ phenotype on SC - leu, trp, ura allowed the growth of microcolonies of the experimental strain, these did not produce any blue color. These results demonstrated that the SC - his, leu, trp, ura media worked very well as a selection for the Ura+ phenotype and that the positive control strain produced strong blue color when grown on this media.

The result of the β–galactosidase assay on SC - trp media indicates that this media may be inappropriate for testing the lacZ reporter in Y166. As shown in Figure 5, both strains grew very robustly on this media. It is clear from the result of the filter β–galactosidase assay that both strains developed a distinct red color (due to the ade2 mutation). The positive control strain is only faintly blue, although it produced a strong blue color on every other media. The experimental strain is faintly blue on this media although it is not blue on any other media. This could be an important consideration when testing if a given "bait" protein activates GAL4-dependent transcription on its own since many of the DNA-binding domain vectors are marked with the *TRP1* gene.

Comparison of the *URA3* and *HIS3* Reporters in Y166

A previous experiment demonstrated that the DBD-TFIIS protein activates the *HIS3* reporter somewhat but does not activate the *URA3* reporter significantly. Since the Y166 strain contains two GAL4-dependent reporters that confer prototrophies, it is possible to select for activation of both reporter constructs simultaneously. One goal of this experiment was to determine if concurrent selection for the His+ phenotype enhanced the selection of the Ura+ phenotype. In addition, in previous experiments I had noticed that the number of cells per plate affected the growth of the experimental strain on media selecting for the Ura+ phenotype. In order to use either the *URA3* or *HIS3* reporter to select for proteins reconstituting GAL4 activity through an interaction with TFIIS, it is necessary that a real positive be distinguishable from the background growth level of a large number of non-positive colonies expressing the DBD-TFIIS protein. This experiment addressed that concern as well.

This experiment investigated the growth of colonies of the positive control strain, Y166 bearing pAS-SNF11 and pNI12, when surrounded by a 100-fold excess of the experimental strain, Y166 bearing pKC14a-1 and pNI12, on media selecting for the Ura+ phenotype or for both the Ura+ and His+ phenotypes. The experimental strain was plated densely to get several thousand cells per plate. On a second set of plates, the same amount of the experimental strain was spiked with approximately 100 cells/plate of the positive control strain, as calculated from the OD_{600}. The actual number was determined by plating an aliquot of the positive control strain separately (data not shown). The number of cells of the experimental strain was visualized by plating on SC - his, leu, trp to select for maintenance of both plasmids but not for reconstitution of GAL4 activity. The Ura+

phenotype was selected on SC - his, leu, trp, ura since it also maintains both plasmid selections and experiments described in the previous section had indicated that this media works well. To select for the His+ phenotype as well as the Ura+ phenotype, 3-AT was added to this media to give final concentrations of 50 mM, 100 mM or 150 mM. All of these plates were incubated at 30°C and were photographed after 5 days, 7 days, and 10 days of growth. The ten day photographs are shown in Figure 7. Filter LacZ assays were performed at 10 days of growth and are shown in Figure 8. For both figures, the A panels show the experimental strain, Y166 bearing pKC14a-1 and pNI12, alone; the B panels show the experimental strain, Y166 bearing pKC14a-1 and pNI12, spiked with the positive control strain, Y166 bearing pAS-SNF1 and pNI12. The type of selective media is indicated below each set of panels.

The abundant growth on SC - his, leu, trp media (panels A-1 and B-1 of Figure 7) shows that the experimental strain was plated densely as intended. On this media the positive control strain does not have a growth advantage over the experimental strain, and its presence is not easily apparent even when assaying for β–galactosidase activity (Figure 8). In contrast, when selecting for the Ura+ phenotype, the positive control cells form large colonies and are vividly blue, while the experimental cells barely grow and produce no blue color (compare A-2 and B-2 panels). Thus it is clear that the growth of cells expressing GAL4 activity reconstituted through two hybrid proteins is easily detectable even in the midst of a large number of cells expressing DBD-TFIIS and an activation domain fusion protein that does not interact.

This experiment also demonstrates that selection for the Ura+ phenotype alone is more effective than selection for both the Ura+ and His+ phenotypes

simultaneously. Cells expressing DBD-TFIIS (and AD-SNF4) do not grow well on any of these media selecting for reconstitution of GAL4 activity (Figure 7, A panels 2-5), nor do the small colonies produce blue color (Figure 8, A panels 2-5). The positive control strain grows significantly better than the experimental strain on all four selective media, although they become quite small at the highest 3-AT concentration. The stimulation of the growth of the experimental strain by 50 mM 3-AT was unexpected. There is almost no growth of these cells on SC - his, leu, trp, ura without 3-aminotriazole, but their growth increased markedly in the presence of 50 mM 3-AT. A concentration of 100 mM 3-AT inhibited the growth of the experimental strain, but the growth of the positive control strain was somewhat inhibited. The optimum selection maximizing the size of colonies of the positive control strain relative to the background growth of the experimental strain was the Ura+ selection alone (compare B-2 with B-3, B-4, and B-5). The absolute correlation between the growth of the positive control colonies and the development of blue color also demonstrated that the lacZ reporter is an effective secondary screen of Ura+ transformants.

Trial Screens of YL2 Library in Y166 Expressing DBD-TFIIS

Several small-scale screens of the YL2 activation domain genomic library (Chien et al., 1991) were conducted to verify that the previously determined selection conditions would be effective when Y166 + pKC14a-1 contained a variety of different AD-fusion proteins instead of only AD-SNF4. The basic transformation procedure is described in the Materials and Methods; a modification to the protocol will be described below. These trials produced useful information and several plasmids encoding proteins putatively interacting with TFIIS.

Comparison of Efficiency Calculations on SC - leu or on SC - his, leu, trp

The first trial was composed of three independent transformations with different amounts of library DNA and the cells were resuspended in SC - his, leu, trp at the end of the transformation procedure. Small aliquots of the cells were plated on either SC - leu or on SC - his, leu, trp in order to determine the transformation efficiency. The first media selects only for receipt of a library plasmid while the second selects for maintenance of pKC14a-1 as well as for receipt of a library plasmid. The reason for the use of both media was to compare transformation efficiencies determined by plating on either of these media. The efficiencies shown in Part 1 of Table 2, section A, calculated from the plating on SC - leu media, are between two and four-fold higher than those shown in Part 2, calculated from the plating on SC - his, leu, trp media. For this reason, all future platings to determine transformation efficiency were done with SC - his, leu, trp media.

Activation of Reporter Genes in Trial Screens

I assayed the colonies from the efficiency platings on SC - leu for β–galactosidase activity. The total number of colonies and the total number of blue colonies, as well as the percentage of colonies which turned blue are displayed in Table 2, section A part 1. This level of positives with the lacZ reporter was far higher than I had anticipated and the result provided a convincing demonstration that the lacZ reporter in Y166 should not be used as the primary screen when DBD-TFIIS is the "bait".

The remainder of each independent transformation was plated on SC - his, leu, trp, ura to test activation of the *URA3* reporter. The percentage of total

colonies which grew on these plates was also quite large (13-20%), see Table 2, section A part 2 and Figure 9. If this large percentage of Ura+ colonies represented activation by DBD-TFIIS, then it would not be possible to use the two-hybrid system to investigate *in vivo* interactions of the TFIIS protein. However, literature from the Elledge laboratory mentioned that when the His+ selection was used, they frequently saw microcolonies. These microcolonies appeared initially but ceased growing and did not produce blue color when assayed for β–galactosidase activity, while real positives continued growing and also produced blue color. I evaluated this possibility as the reason for the large fraction of colonies which had grown on the selective media for the Ura+ phenotype in the transformation shown in Figure 9. I wanted to determine if these putatively Ura+ colonies also produced β–galactosidase activity.

To distinguish between microcolonies which appeared initially on media lacking uracil and colonies that would continue to grow, the cells from this transformation were replica plated (13 days after the initial plating) onto fresh SC - his, leu, trp, ura plates and incubated at 30°C for 5 days. The colonies were then filter lifted and assayed for β–galactosidase activity. Only one of the colonies from all three of the transformations produced blue color (Figure 10). This colony had been growing noticeably better than the other colonies prior to replica plating and became more pronounced afterwards. This colony was restreaked on SC - his, leu, trp, ura. The library plasmid contained in this putative positive was designated pYL2-1, the first isolate from the YL2 library.

A second trial transformation was performed in essentially the same manner except the cells were replica plated to new SC - his, leu, trp, ura plates after only 7 days of growth at 30°C. Younger colonies not only replica plate more efficiently,

they also assay for β-galactosidase more reproducibly. The new plates were incubated at 30°C for 4 days. At this time, the colonies were lifted for β-galactosidase assays. However, the majority of the colonies were still quite small and the results were inconclusive.

These replica plates were allowed to continue growing at room temperature for 14 days, at which time a few colonies had become very large while the majority remained quite small. The largest colony was restreaked onto fresh media and named pYL2-2. Three other colonies of an intermediate size were restreaked and named pYL2-3, pYL2-4, and pYL2-5. Colonies from these restreaks were lifted for β-galactosidase assays. The results were not as convincing as the result with pYL2-1 (Figure 10). However, each of the four colonies picked on the basis of size developed blue tinged edges (data not shown). In addition, there were two other colonies with faint blue edges. These were restreaked on fresh media and named pYL2-6 and pYL2-7.

The percentage of Ura+ colonies in the second trial was a sufficiently low percentage of the total number of colonies to proceed with characterization of the putative positives identified in the trial screens (Table 2, section B). However, in the future, a simple modification of the transformation protocol might improve the initial selection for Ura+ colonies. The transformation protocol had been intended for use of the *HIS3* reporter and at the end of the procedure the cells were resuspended in SC - his, leu, trp. Although this media is not selective for the *HIS3* reporter as it lacks 3-AT, it does minimize the amount of histidine in the cells. For the next trial screen, I used SC - his, leu, trp, ura for the resuspension media to minimize the amount of uracil present in the cells. When these cells were plated on the selective media SC - his, leu, trp, ura, the Y166 cells containing pKC14a-1

and which received a library plasmid did not grow. In contrast, colonies of the positive control strain grew well and produced vivid blue color (data not shown). With this modification of the protocol described in the Methods, it should be possible to screen an activation domain library using DBD-TFIIS as the "bait".

Further Characterization of Putatively Positive pYL2 Plasmids

The trial screens had isolated several plasmids that produced a Ura+ phenotype in combination with pKC14a-1. These were investigated further to provide greater confidence that the DBD-TFIIS protein could be used as the "bait" in a two-hybrid screen. The pYL2-1 isolate had been restreaked several times and the different streaks did not display identical growth, so pYL2-1 was split into pYL2-1A, pYL2-1B, and pYL2-1C. Each was restreaked on fresh media at the same time as the six isolates selected from the second trial screen. The cells were incubated at 30°C for 10 days (Figure 11) and then assayed for β–galactosidase activity (Figure 12). Cells containing any isolate of pYL2-1, pYL2-2, or pYL2-4, grew well. Cells containing pYL2-2 along with pKC14a-1 grew especially well. Cells containing any of the other four plasmids, pYL2-3, pYL2-5, pYL2-6, or pYL2-7, grew poorly, especially those containing either pYL2-6 or pYL2-7. The intensity of blue color was correlated with the growth. The colonies with pYL2-2 produced an intense blue color and those with pYL2-1 and pYL2-4 produced strong blue color. Colonies containing any of the other four plasmids did not produce convincing blue color.

Further characterization required the isolation of these pYL2 plasmids. The putatively positive plasmids were rescued into *E. coli* as described in the Materials and Methods. For each of the plasmids (pYL2-1A and pYL2-1B were both

characterized, but not pYL2-1C), two *E. coli* colonies were restreaked on LB - ampicillin plates. DNA was prepared from both isolates of each pYL2 plasmid and transformed back into Y166 cells containing pKC14a-1, pAS-SNF1, or no plasmid. Transformants containing the library plasmids were selected on either SC - his, leu, trp, or on SC - leu, as appropriate. The transformants were filter lifted for β–galactosidase assays. Both isolates of all eight plasmids produced blue color when reintroduced into Y166 bearing pKC14a-1 (data not shown). None of these plasmids stimulated production of blue color when transformed into Y166 or into Y166 bearing pAS-SNF1. These assays indicated that none of these plasmids are capable of stimulating transcription of the reporter constructs, either alone or in the presence of a different GAL4 DNA-binding domain fusion.

The sizes and restriction patterns of the inserts in each of the plasmids isolated from the YL2 library was determined by restriction analysis with the enzyme HindIII (Figure 13). Digestion of the plasmids pYL2-3, pYL2-5, pYL2-6, and pYL2-7 with this enzyme produced the expected digestion pattern of library vector. The restriction analysis indicated that pYL2-1A and pYL2-1B are not identical. This difference was likely produced by a recombination event in Y166 cells and did not alter the initial portion of the protein fused to GAL4-AD, based on subsequent sequence analysis. The three plasmids, pYL2-1, pYL2-2, and pYL2-4, which retested positively for the Ura+ phenotype and for blue color (Figures 11 and 12) contain inserts of different sizes and restriction patterns.

To determine the identities of the proteins encoded by the three pYL2 plasmids containing inserts, each was sequenced using a primer that anneals within the GAL4 activation domain sequence. Sequencing confirmed that pYL2-3, pYL2-5, and pYL2-6 are identical to pGAD2 and indicated that pYL2-7 contains a

43 base pair fragment of the POL1 gene in the wrong orientation. Thus, all four of these, which conferred inefficient selection and low β-galactosidase activity, would not be considered positives. The protein encoded by pYL2-4 is not represented in the databases. The plasmid pYL2-2 encodes a portion of the MSN4 protein, one of two zinc-finger proteins identified as multicopy suppressers of a defect in the SNF1 protein kinase (Estruch and Carlson, 1993). The carboxyl-terminal region containing both of the two zinc finger motifs is present in pYL2-2. Sequencing of pYL2-1A and pYL2-1B produced identical results and indicates that the pYL2-1 plasmid encodes a previously unidentified yeast protein with homology to the Smcy and Smcx (Xe169) proteins of mice and humans (Agulnik et al., 1994a, 1994b; Wu et al., 1994a, 1994b) and to the human retinoblastoma binding protein 2 (RBP2; Figure 14; Defeo-Jones et al., 1991; Fattaey et al., 1993).

DISCUSSION

In this work, I have demonstrated that it should be possible to use the two-hybrid technique to identify proteins which interact with the transcription elongation factor TFIIS *in vivo*, even though this factor interacts with RNA polymerase II *in vitro*. Testing the ability of a fusion of the TFIIS protein to the DNA-binding domain of the *GAL4* protein to activate GAL4-dependent reporter genes indicated that the DBD-TFIIS fusion protein produces a slightly higher level of activation of GAL4-dependent reporters than does the DBD-SNF1 fusion protein. In the Y153 strain, the level of activation prohibits use of the lacZ reporter. The Y166 strain contains a *URA3* reporter which requires a higher level of activation than other reporter constructs. As it is not activated by DBD-TFIIS, it should be possible to screen an activation domain library for proteins interacting with TFIIS by selection for the Ura+ phenotype in this strain. The lacZ and *HIS3* reporters behave differently in Y166 than in Y153, perhaps due to different locations of integration or different promoter constructions. The low level of activation of either of these reporters by DBD-TFIIS suggests that either is usable as a secondary criterion to confirm Ura+ selectants.

Although the background level of growth in the Ura+ selection was somewhat high in the two trial screens described, I think that a change in resuspension media to minimize the amount of uracil present in the cells when they are plated on the selective media will correct this problem, as described above. Modifications to decrease the level of expression of the DBD-TFIIS fusion protein might also improve the selection (Fields and Sternglanz, 1994); one possibility would be to move the DBD-TFIIS protein to a single copy vector. Based

on the results described above, I feel it will be possible to screen an activation domain library to identify proteins which interact with TFIIS *in vivo*.

The trial screens isolated several putative positives from the YL2 library. Although there was a problem with a high level of background growth in the selection for Ura+ colonies in these two trial screens, two of the putative positives, pYL2-1 and pYL2-2, behaved in the manner appropriate for encoding a protein that interacts with TFIIS. During the secondary Ura+ selection, the colonies containing either of these plasmids grew to a much larger size than the surrounding colonies, which did not continue to grow. The next three plasmids, pYL2-3, pYL2-4, and pYL2-5 were selected because they were somewhat larger than the surrounding colonies. The last two pYL2-6 and pYL2-7 were selected based on faint blue color in the lacZ assay.

Four of these last five plasmids did not retest positively for activation of the *URA3* reporter (Figure 11) and the small colonies on this media did not produce strong blue color (Figure 12). Three of these plasmids proved to be the library vector without an insert and the fourth contained a peptide insert in the wrong orientation. These plasmids did produce blue color when isolated and transformed back into Y166 bearing pKC14a-1. However, these assays were not done on media selecting for the Ura+ phenotype. This suggests that the Ura+ selection should be the primary criterion, with the lacZ used only to retest Ura+ selectants.

Three of the plasmids, pYL2-1, pYL2-2, and pYL2-4, retested positively when restreaked on media selecting for the Ura+ phenotype (Figure 11) and these colonies produced vivid blue color, especially those containing pYL2-4 (Figure 12). All three restored blue color when transformed back into Y166 expressing DBD-TFIIS, although this should be repeated on the appropriate media as described

above. None of these plasmids stimulated activation of the lacZ reporter when isolated and transformed into Y166 without a GAL4 DBD fusion protein or into Y166 expressing the DBD-SNF1 protein. This certainly indicates that these activation domain fusion proteins are insufficient to stimulate GAL4-dependent transcription alone or in the presence of a different DBD fusion protein. Sequencing indicated that pYL2-1 encodes a previously unknown yeast protein with homology to several mammalian proteins, that pYL2-2 encodes a portion of the yeast MSN4 protein, and that pYL2-4 encodes a protein not currently represented in the databases.

The determination that pYL2-1 encodes a yeast protein with significant homology to the Smcy and Smcx (Xe169) proteins and to retinoblastoma binding protein 2 (Figure 14) is interesting. However, the available information on these proteins does not suggest much about how this protein might interact with TFIIS. The murine and human Smcy and Smcx proteins are located on the Y and X chromosomes respectively. Smcy is widely transcribed in male tissues and is a source of H-Y antigen (Scott et al., 1995; Wang et al., 1995). Smcx escapes X-inactivation in females (Agulnik et al., 1994a; 1994b; Wu et al., 1994a, 1994b). Recently, a mouse protein, *jumonji*, with homology to Smcx (Xe169) has been implicated as required for neural tube formation (Takeuchi et al., 1995). RBP2 was originally cloned by screening a cDNA expression library for proteins which interacted with a GST fusion to the retinoblastoma protein (Rb) and only partial cDNAs were cloned (Defeo-Jones et al., 1991); demonstration of an *in vivo* interaction with Rb has been more difficult (Fattaey et al., 1993, P. Whyte, personal communication). Furthermore, the region of RBP2 with homology to the protein encoded by pYL2-1 shown in Figure 14 is not present in an RBP2 construct

demonstrated to bind to Rb *in vitro* (Kim et al., 1994b). The sequence in pYL2-1 is similar to the region of RBP2 with homology to Smcy and Smcx and which contains zinc-finger and homeobox motifs (Scott et al., 1995, Fattaey et al., 1993). RPB2 has been proposed to be a transcription factor. Further sequencing of pYL2-1 would indicate the size of this yeast protein and whether the homology to any of the mammalian proteins extends further.

The identification of the MSN4 protein as putatively interacting with the TFIIS protein by activation of GAL4-dependent reporter genes in a two-hybrid screen of an activation domain library is interesting since MSN4 and the related protein MSN2 were identified by their effects on carbon utilization (Estruch and Carlson, 1993). The carboxyl-terminal regions of MSN4 and MSN2 contain two zinc finger motifs of the Cys_2His_2 type with homology to the yeast proteins MIG1 and RGM1 and to the mammalian early growth response and Wilms' tumor locus zinc finger proteins. Deletion of both MSN2 and MSN4 causes defects in carbon utilization, including galactose. The zinc finger regions bind DNA and overexpression of this region of either protein is deleterious to growth. Fusions of either protein to the LexA DNA-binding domain functioned as strong transcriptional activators (Estruch and Carlson, 1993). Thus MSN4 itself might affect the transcription of GAL4-dependent promoters. However, the fusion of the zinc finger region of MSN4 to the GAL4 activation domain did not activate transcription of a GAL4-dependent reporter, even when the GAL4 DNA-binding domain fusion to the SNF1 protein was also present.

Further work will be required to determine if these proteins actually interact with TFIIS and whether an interaction is meaningful. If these interactions are real, then these proteins will be isolated again in subsequent screens of this activation

domain library which incorporate the suggested improvements. I have speculated that the TFIIS protein might play a role in sensing or responding to nutrient levels based on the changes in flocculence when the *PPR2* gene is deleted or overexpressed (K. Christie and K. Johnson, unpublished observations). Although the occurrence of flocculence is not well understood, it is observed that it is stimulated by some nutrient conditions and not by others (M. Stratford, 1992a; M. Stratford, 1992b; Mota and Soares, 1992; Smit et al., 1992). In addition, although it is clear that the deletion of the *PPR2* gene does not produce a growth defect during logarithmic growth, unless 6-azauracil is present, cells lacking the *PPR2* gene seem to be defective in either entry or exit from stationary phase (K. Christie, K. Johnson, and J. Davie, unpublished observations). It has been suggested that the interaction of TFIIS with RNA polymerase II is important for the polymerase to elongate through pause sites *in vivo* and that this step can become rate limiting if the polymerase is defective (Archambault et al., 1992b) or when nucleotide pools are limiting (Exinger and Lacroute, 1992). Thus TFIIS function may be more important under conditions where nucleotides are less abundant, such as nutrient limitation or stationary phase (Werner-Washburne et al., 1993). Use of the two-hybrid system to identify proteins which interact with TFIIS *in vivo* may begin to elucidate the role of this transcription elongation factor in cells.

TABLE LEGENDS

Table 1 - List of Strains

The name and relevant genotype of each strain of *Escherichia coli* and *Saccharomyces cerevisiae* used is given. The *E. coli* strains DH10B (Grant et al., 1990) and XL-1Blue (Bullock et al., 1987) were obtained from Stratagene. MH1 (Hall et al., 1984) was obtained from K. Arndt (Univ. of Pittsburgh). KC8 was obtained from J. Rine. *S. cerevisiae* strains, Y153 and Y166, were obtained from S. Elledge.

Table 1 - List of Strains

Escherichia coli	Genotype
XL-1Blue	*recA1 endA1 gyrA96 thi hsdRIT($r^-_K m^+_K$) SupE44 relA1 λ^- (lac) {F' proAB lac1^q ZΔM15 Tn10(tet^R)*
KC8	*hsdR$^-$ (hsdM$^+$) leuB600 trpC9830 pyrF::Tn(Kan^R) galU galK hisB463 ΔlacX74 strA*
MH1	*araD139 lacX74 galU galK hsr rpsL*
DH10B	F$^-$ *araD139 Δ(ara, leu)*7697 *ΔlacX74 galU galK rpsL deoR* Φ80d*lacZ*ΔM15 *endA1 nupG recA1 mcrA Δ(mrr hsdRMS mcrBC)*

Saccharomyces cerevisiae	Genotype
Y153	MAT**a** ura3-52 leu2-3,-112 his3-Δ200 ade2-101 trp1-901 gal4Δ gal80Δ LYS2::GAL1-->HIS3 URA3::GAL1-->lacZ
Y166	MAT**a** ura3-52 leu2-3,-112 his3-Δ200 ade2-101 trp1-901 gal4Δ gal80Δ LYS2::GAL UAS-->HIS3 RNR::GAL1-->URA3 GAL1-->lacZ (unknown location)

Table 2 - Results of Preliminary Activation Domain Library Screens

Relevant information from two sets of transformations of activation libraries into a Y166 bearing pKC14a-1 are given. Section A gives information on the first transformation of YL2 library DNA into Y166 + pKC14a-1. Part 1 gives data for assays of the lacZ reporter. These transformation efficiencies were calculated from the plating on SC - leu media. Part 2 gives data for the Ura+ selection before replica plating. These transformation efficiencies were calculated from the plating on SC - his, leu, trp media. Section B gives information for the Ura+ selection after replica plating. In all cases, Theoretical Number of Colonies is based on the transformation efficiency shown in the column to the left and represents the estimated number of transformants in the entire transformation mixture. For B, the minimum estimate (determined by counting) of the number of colonies growing on selective media for Ura+ phenotype prior to replica plating was 250 colonies, which would be 10% of the theoretical total number of colonies.

Table 2 - Results of Three Preliminary Activation Domain Library Screens

A - Transformations of YL2 Library DNA into Y166 + pKC14a-1, #1				
1. Screen for Blue Color on SC - leu				
Transformation	Calculated Transformation Efficiency Colonies/ug	Total # of Colonies	Total # of Blue Colonies	**% of Colonies Blue in LacZ Assay**
# 1	1963	123	22	**17.9**
# 2	2864	398	72	**18.1**
# 3	3584	564	90	**16.0**
2. Selection for Ura+ Phenotype on SC - his, leu, trp, ura				
Transformation	Calculated Transformation Efficiency Colonies/ug	Theoretical # of Colonies	Total # of Colonies on SC - his, leu, trp, ura before replica plating	**% of Colonies Growing on Ura+ Selection**
# 1	755	1125	230	**20.4**
# 2	982	2925	620	**21.2**
# 3	866	3870	515	**13.3**
B - Transformation of YL2 Library DNA into Y166 + pKC14a-1, #2				
Selection for Ura+ Phenotype				
Transformation	Calculated Transformation Efficiency Colonies/ug	Theoretical # of Colonies	# Growing well on SC - his, leu, trp, ura after replica plating	**% of Colonies Growing on Ura+ Selection**
# 1	831	2475	6	**0.24**

FIGURE LEGENDS

Figure 1 - Diagram of Scheme for Two Hybrid Screen

The basic scheme for a two hybrid screen is diagrammed here. **A.** The pKC14a-1 plasmid expresses a fusion of the GAL4 DNA-binding domain to the yeast TFIIS protein, DBD-TFIIS. The DBD-TFIIS binds to its recognition sites within the GAL4-dependent promoter but does not stimulate transcription. In order to proceed to screening an activation domain library, it is crucial that this fusion protein does not activate transcription of the integrated reporter construct driven by a GAL4-dependent promoter. **B.** The activation domain library is constructed such that each library protein is fused to the activation domain of the GAL4 protein. By itself, the activation domain does not bind to the GAL4-dependent promoter. Thus the activation domain-library protein fusions should not stimulate transcription of the integrated reporter construct. **C.** If a library protein interacts with the TFIIS protein, tethered to the GAL4-dependent promoter by the DNA-binding domain, then this interaction will bring the GAL4 activation domain into the proximity of promoter and transcription of the reporter construct will be activated.

Scheme for Using a Dihybrid Screen to Identify Proteins
Which Interact with Yeast TFIIS

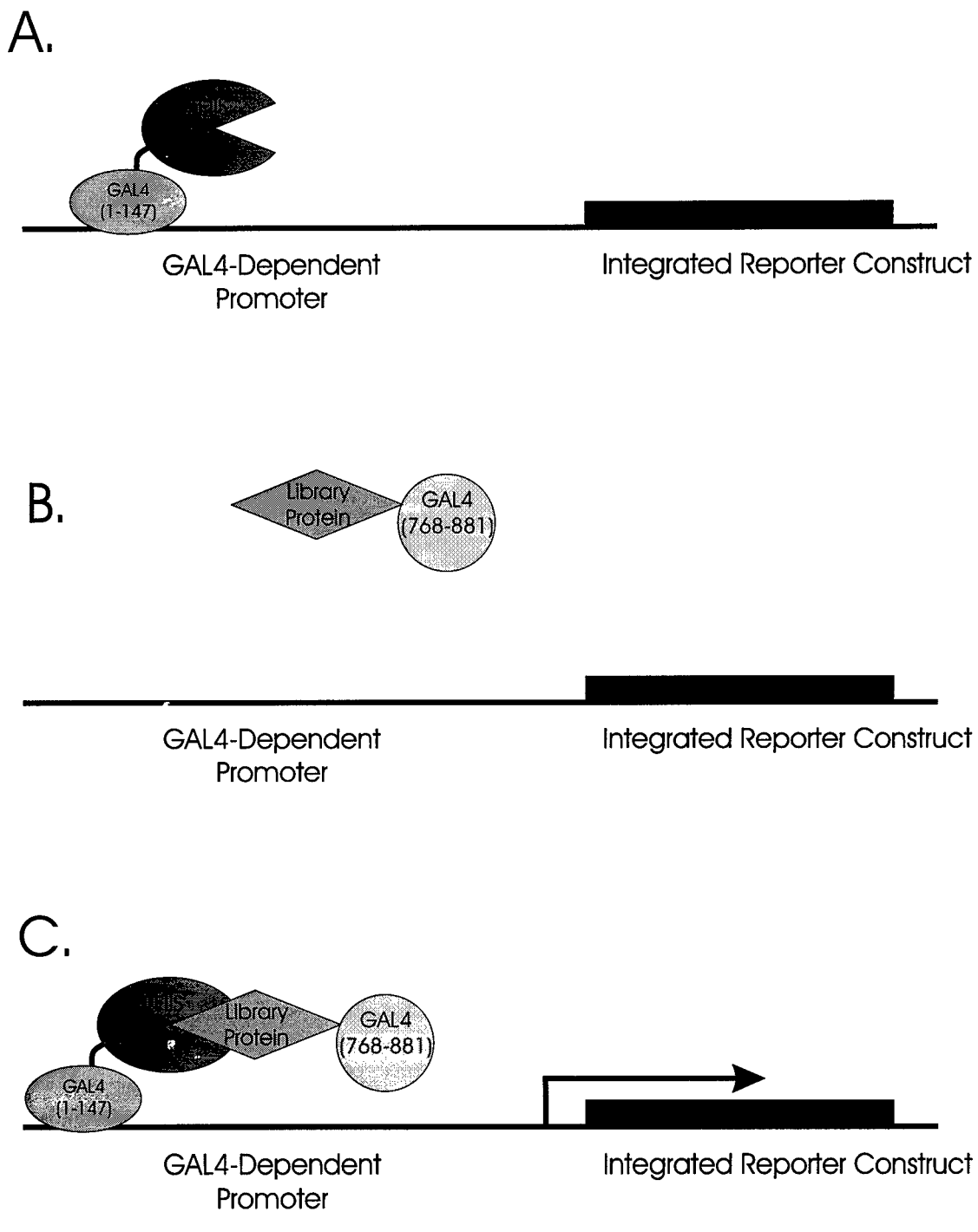
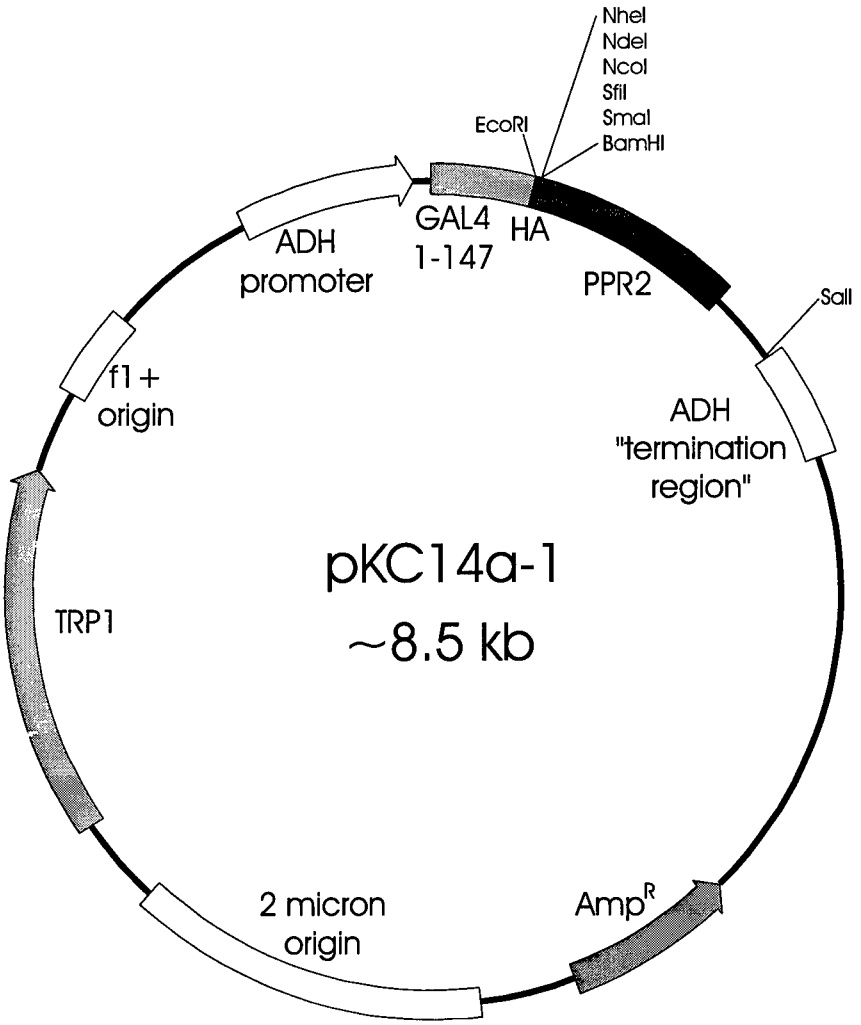


Figure 2 - Diagram of pKC14a-1

A. The map of the plasmid pKC14a-1 is shown here. Some restriction enzyme sites are indicated. **B.** The sequence of the junction of the HA epitope and polylinker of pAS1 with the open reading frame of *PPR2* is shown. The reading frame is indicated by the triplet spacing. Relevant restriction enzyme sites are indicated with bars above or below the recognition site.

A.



B.

GAA TCC ATG GCT TAC CCA TAC GAT GTT CCA GAT TAC GCT AGC TTG GGT

EcoRI — NheI

GGT CAT ATG GCC ATG GAG GCC CCG GGG ATC CCG ATG GAT AGT AAG

SfiI — BamHI

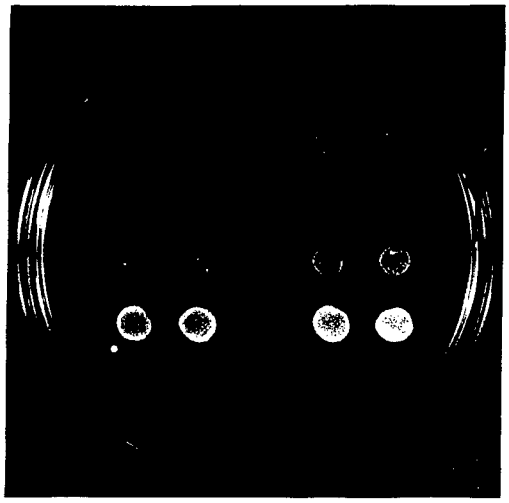
NdeI — NcoI — SmaI — PPR2 ORF

Figure 3 - Comparison of *HIS3* and *URA3* reporters of Y166 with *HIS3* Reporter of Y153

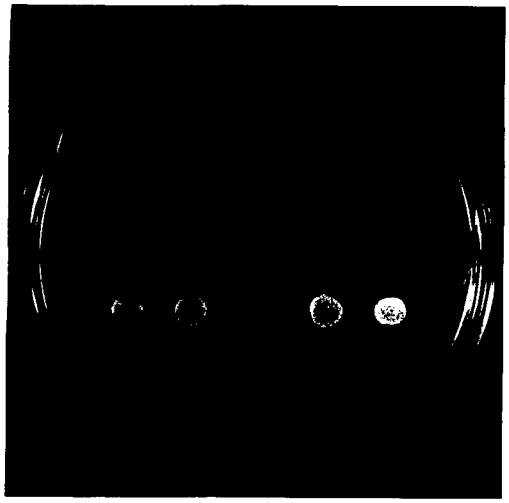
3-aminotriazole (final concentration indicated below each panel) was spread on solid minimal media lacking histidine and trpytophan (SC - his, trp) and allowed to soak in until dry. Colonies of test strains were picked into sterile water in wells and replica plated onto each concentration of 3-aminotriazole and also onto solid minimal media lacking uracil (SC - ura). The following chart indicates the strain and the plasmid(s) it bears present in each well. The pGAL4-laminC plasmid encodes a fusion of the GAL4-DBD to the human lamin C protein; since it is marked with the LEU2 gene, it does not allow growth on media lacking tryptophan.

Y153 + pKC14a-1	Y153 + pKC14a-1	-	Y166 + pKC14a-1	Y166 + pKC14a-1
Y153 + pAS-SNF1	Y153 + pAS-SNF1	Y153 + pCL1	Y166 + pAS-SNF1	Y166 + pAS-SNF1
Y153 + pKC14a-1, pNI12	Y153 + pKC14a-1, pNI12	Y153 + pCL1	Y166 + pKC14a-1, pNI12	Y166 + pKC14a-1, pNI12
Y153 + pAS-SNF1, pNI12	Y153 + pAS-SNF1, pNI12	Y166 + pCL1	Y166 + pAS-SNF1, pNI12	Y166 + pAS-SNF1, pNI12
Y153 + pGAL4-laminC	Y153 + pGAL4-laminC	Y166 + pCL1	Y166 + pGAL4-laminC	Y166 + pGAL4-laminC

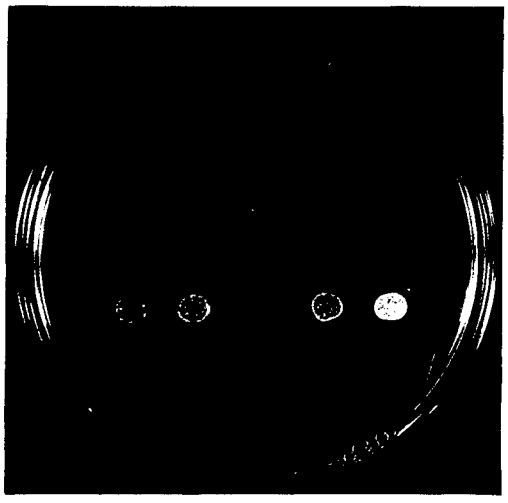
Comparison of HIS3 and URA3 Reporters in Y166
with the HIS3 Reporter in Y153



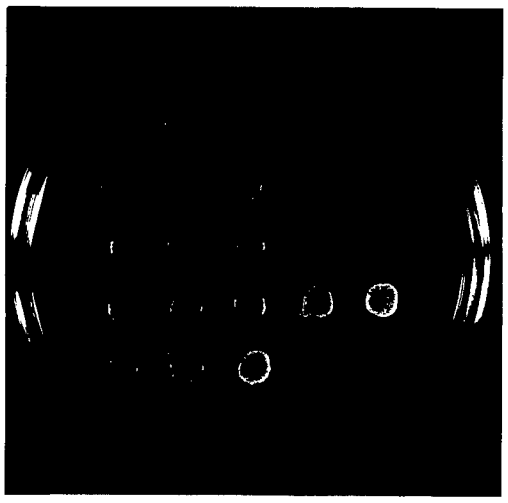
SC - his, trp + 25 mM 3-AT



SC - his, trp + 50 mM 3-AT



SC - his, trp + 75 mM 3-AT

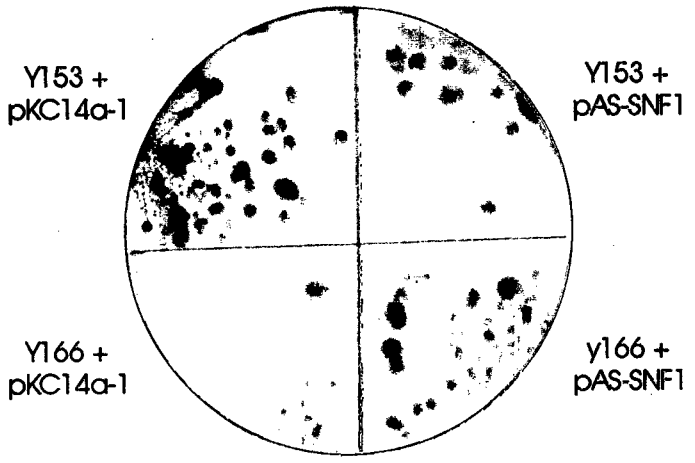


SC - ura

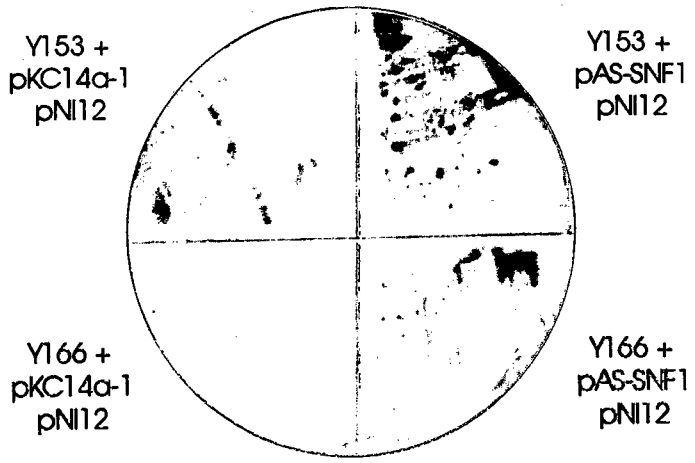
Figure 4 - Comparison of LacZ Reporters in Y153 and Y166

The test strains indicated were streaked on the media indicated and incubated at 30°C. When colonies had grown large enough to lift, the cells were lifted onto 7 cm Whatman filters and assayed for β-galactosidase activity as described in the Methods.

Comparison of Background for LacZ Reporters in Y153 and in Y166



SC - trp



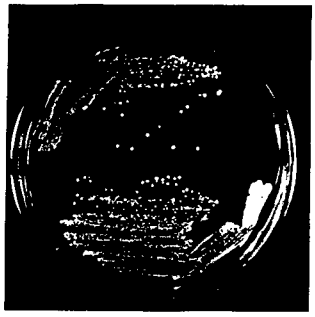
SC - his, leu, trp

Figure 5 - Growth of Y166 + pKC14a-1, pNI12 and Y166 + pAS-SNF1, pNI12 on Various Selective Media

The positive control strain Y166 + pAS-SNF1, pNI12 and the experimental strain Y166 + pKC14a-1, pNI12 were streaked on the media indicated below each panel. The positive control strain is on the top half of each plate; the experimental strain is on the lower half of each. The cells were incubated at 30°C and photographed after 5 days, 7 days, and 10 days of growth. The photographs after 10 days growth are shown here.

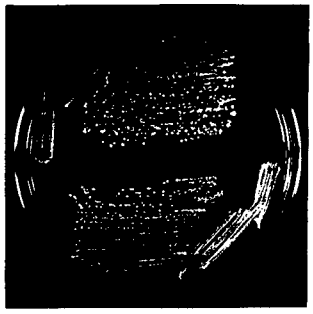
Growth of Y166 + pKC14a-1, pNI12 and Y166 + pAS-SNF1, pNI12
on Various Selective Media

D.



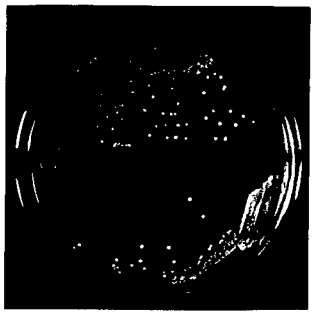
SC - leu, trp

A.



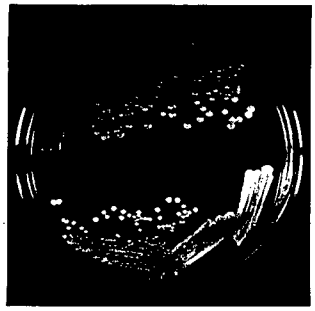
SC - leu

E.



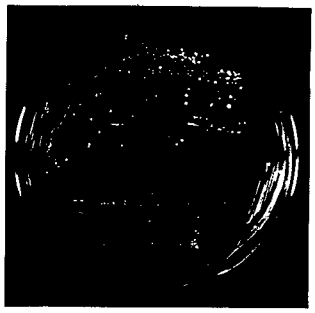
SC - his, leu, trp

B.



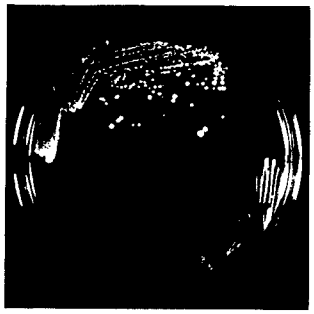
SC - trp

F.



SC - leu, trp, ura

C.



SC - ura

G.

SC - his, leu, trp, ura

Figure 6 - Assays for β–Galactosidase Activity of Y166 + pKC14a-1, pNI12 and Y166 + pAS-SNF1, pNI12 on Various Selective Media

After 10 days of growth, the cells on the plates shown in Figure 5 were lifted on 7 cm Whatman filters and assayed for β–galactosidase activity as described in the methods. As in Figure 5, the positive control strain is on the top half of each plate; the experimental strain is on the lower half of each.

Assays for β-Galactosidase Activity of Y166 + pKC14a-1, pNI12 and Y166 + pAS-SNF1, pNI12 on Various Selective Media

D.

SC - leu, trp

A.

E.

SC - leu

SC - his, leu, trp

B.

F.

SC - trp

SC - leu, trp, ura

C.

G.

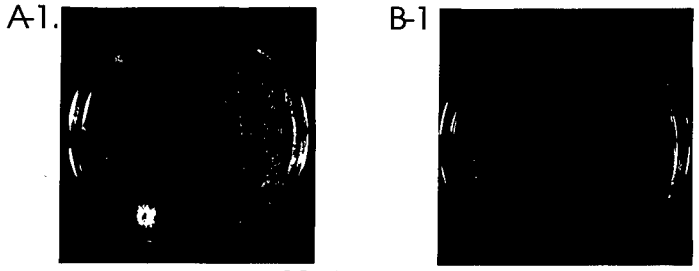
SC - ura

SC - his, leu, trp, ura

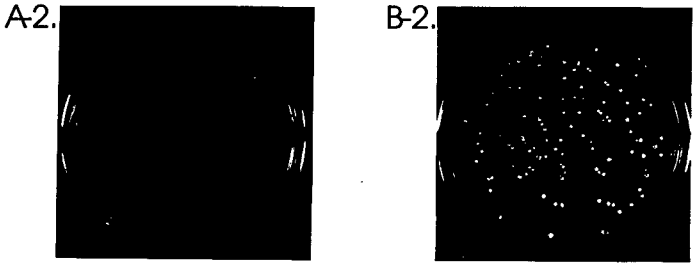
Figure 7 - Growth of Y166 + pKC14a-1, pNI12 or
Y166 + pKC14a-1, pNI12 and Y166 + pKC14a-1, pNI12
on Various Selective Media

Two cultures were grown up, one of Y166 + pKC14a-1, pNI12 and one of Y166 + pAS-SNF1, pNI12. The OD_{600} of each was determined. Both cultures were grown to a similar density, an OD of 1.7 for the experimental strain and of 2.0 for the positive control strain. **A.** A 100-fold dilution of Y166 + pKC14a-1, pNI12 was prepared and one tenth of it was plated on each of the five types of media, as indicated below. **B.** A separate 100-fold dilution of Y166 + pKC14a-1, pNI12 was spiked with Y166 + pAS-SNF1, pNI12 (dilution of the control strain in the mix was 10,000-fold). One tenth of this mix was plated on each of the five media. All plates were incubated at 30°C and were photographed after 5 days, 7 days, and 10 days of growth. Growth after 10 days is shown here.

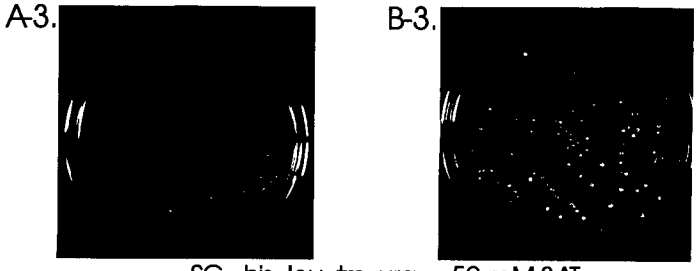
Growth of Y166 + pKC14a-1, pNI12
or Y166 + pKC14a-1, pNI12 and Y166 + pAS-SNF1, pNI12
on Various Selective Media



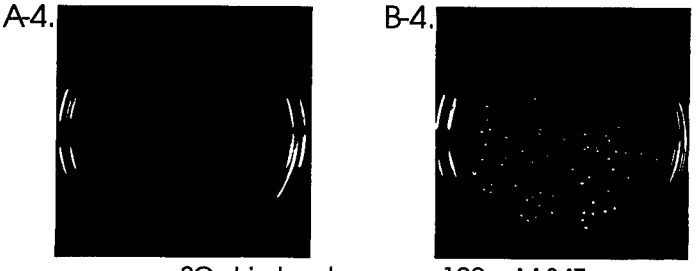
SC - his, leu, trp



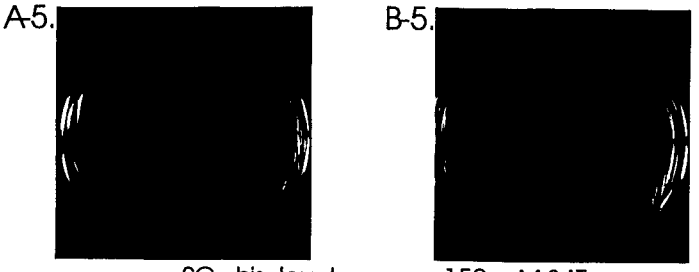
SC - his, leu, trp, ura



SC - his, leu, trp, ura + 50 mM 3AT



SC - his, leu, trp, ura + 100 mM 3AT



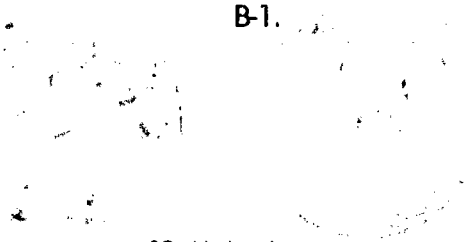
SC - his, leu, trp, ura + 150 mM 3AT

Figure 8 - LacZ Assays of Y166 + pKC14a-1, pNI12 or Y166 + pKC14a-1, pNI12 and Y166 + pKC14a-1, pNI12 on Various Selective media

After 10 days of growth, the cells on the plates shown in Figure 7 were lifted on 7 cm Whatman filters and assayed for β–galactosidase activity as described in the Methods.

Assays for β-Galactosidase Activity of Y166 + pKC14α-1, pNI12 or Y166 + pKC14α-1, pNI12 and Y166 + pAS-SNF1, pNI12 on Various Selective Media

A-1. B-1.



SC - his, leu, trp

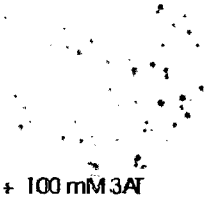
A-2. B-2.

SC - his, leu, trp, ura

A-3. B-3.

SC - his, leu, trp, ura + 50 mM 3AT

A-4. B-4.



SC - his, leu, trp, ura + 100 mM 3AT

A-5. B-5.

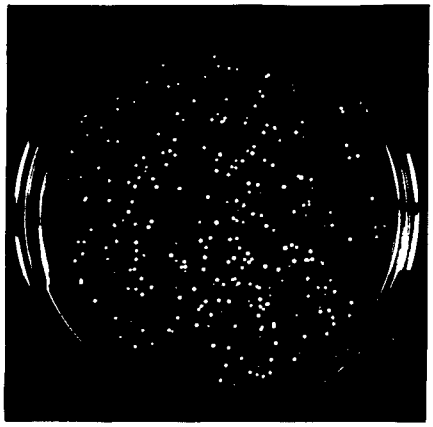
SC - his, leu, trp, ura + 150 mM 3AT

Figure 9 - Y166 + pKC14a-1 Transformed with YL2 Library DNA
10 Days Growth on SC - his, leu, trp, ura

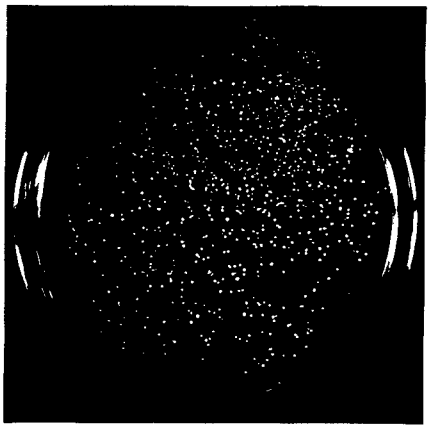
The first set of trial transformations of YL2 library DNA into Y166 + pKC14a-1 was plated on solid minimal media lacking histidine, leucine, tryptophan, and uracil (SC - his, leu, trp, ura) as described in the Methods and Results and incubated at 30°C. This set was composed of three independent transformations. #3 was split between two plates selective for the Ura+ phenotype. Growth after 10 days is shown here.

Y166 + pKC14a-1 Transformed with YL2 Library DNA
10 Days Growth on SC - his, leu, trp, ura

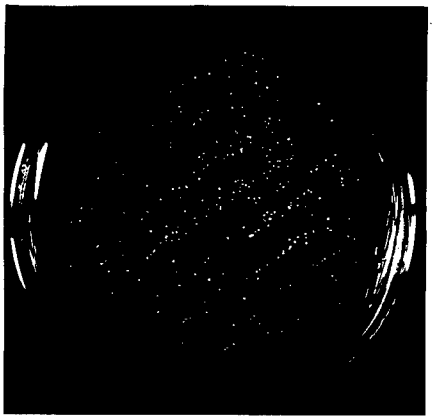
Transformation # 1



Transformation # 2



Transformation # 3 - plate A



Transformation # 3 - plate B

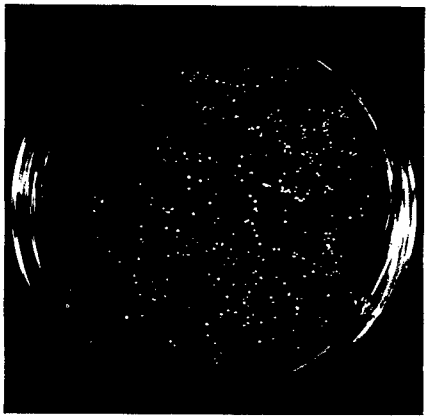


Figure 10 - β–galactosidase Assays of Transformations and Control Strains

After 13 days on SC - his, leu, trp, ura plates, the cells on the four plates shown in Figure 9 were replica plated to fresh SC - his, leu, trp, ura plates and incubation at 30°C continued for 5 more days. At this time the cells were lifted on 7 cm Whatman filters and assayed for β–galactosidase activity as described in the Methods.

Assays for β-Galactosidase Activity of Transformations and Control Strains

Y166 + pAS-SNF1, pNI12

Y166 + pKC14a-1, pNI12

Transformation # 1

Transformation # 2

Transformation # 3 - plate A

Transformation # 3 - plate B

Figure 11 - Growth of Putatively Positive pYL2 Plasmids

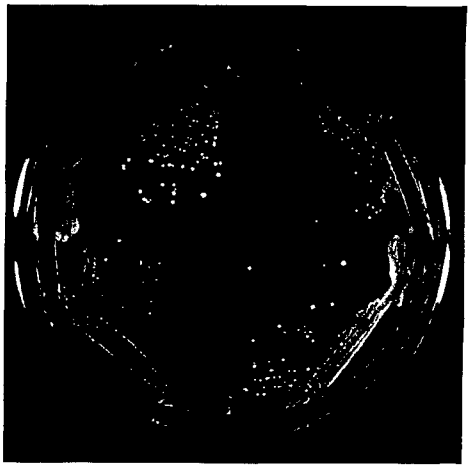
All of the putatively positive pYL2 plasmids, including all three variants of pYL2-1, isolated from the YL2 activation domain library were streaked on SC - his, leu, trp, ura and incubated at 30°C for 10 days. Growth after 10 days is shown here.

Growth of Putatively Positive pYL2 Plamids

pYL2-1A pYL2-1C

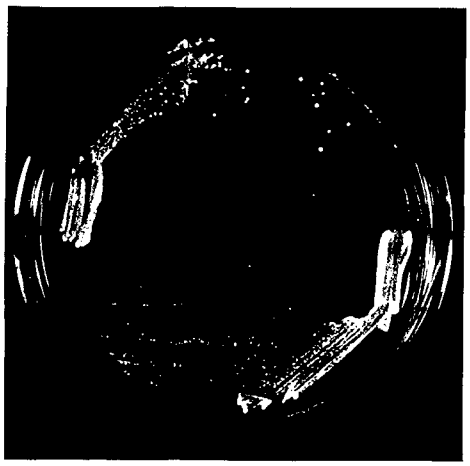
pYL2-1B

pYL2-2 pYL2-5



pYL2-3 pYL2-4

pYL2-6



pYL2-7

Figure 12 - β-Galactosidase Assays of Cells Containing Putatively Positive pYL2 Plasmids

After 10 days of growth, the cells on the plates shown in Figure 11 were lifted on 7 cm Whatman filters and assayed for β–galactosidase activity as described in the Methods.

β-Galactosidase Activity of Putatively Positive pYL2 Plamids

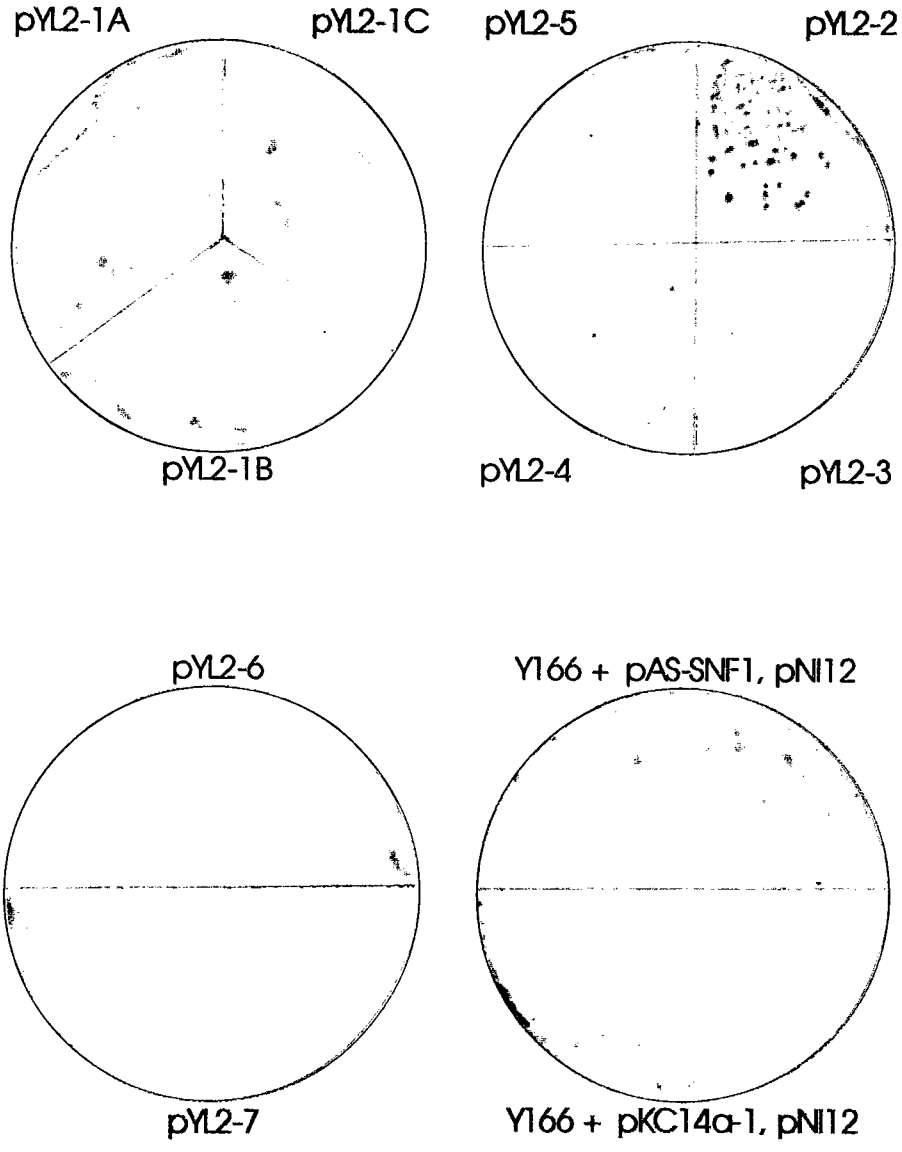


Figure 13 - Restriction Analysis of Putatively Positive pYL2 Plasmids

The YL2 library was constructed in the pGAD2 vector (Chien et al., 1991). Digestion of pGAD2 with the enzyme HindIII produces the vector fragment of 12 kb and a 700 base pair fragment containing the insertion site for the library fragments. Each of the putatively positive plasmids from the trial screens was digested with HindIII and the sizes of the resulting fragments are shown. Digestion of the plasmids pYL2-3, pYL2-5, pYL2-6, and pYL2-7 with this enzyme produced two fragments consistent with the expected digestion pattern of pGAD2. The restriction analysis indicated that pYL2-1A and pYL2-1B are not identical. The digestion of pYL2-1A produced five fragments, while digestion of pYL2-1B produced only two. Digestion of pYL2-2 produced three fragments; digestion of pYL2-4 produced four fragments. Thus the three plasmids, pYL2-1, pYL2-2, and pYL2-4, which retested positively for the Ura+ phenotype and for blue color (Figures 11 and 12) contain inserts of different sizes and restriction patterns.

Figure 13 - Restriction Analysis of Putatively Positive pYL2 Plasmids

pGAD2 Vector	HindIII Fragments of Plasmids (size in base pairs)							
	pYL2-1A	pYL2-1B	pYL2-2	pYL2-3	pYL2-4	pYL2-5	pYL2-6	pYL2-7
12000	12000	12000	12000	12000	12000	12000	12000	12000
700	5000	5000	4500	700	4500	700	700	700
	3500		3900		3600			
	2500				1000			
	1300							

Figure 14 - Alignment of Predicted Amino Acid Sequence from pYL2-1B

The predicted amino acid sequence of the protein fused to the activation domain of the *GAL4* protein in the plasmid pYL2-1B is shown in one letter code above the dividing line. Lowercase x's indicate codons which could not be predicted because the nucleic acid sequence contained an ambiguity. The alignments of the first five sequences identified by the blastp program (Altschul et al., 1990) are shown below the divider. Blastp identified three regions of homology with each of this group of proteins. In these five sequences, the bold letters indicate identical residues relative to pYL2-1B; the underlined letters indicated conserved residues. The bold letters in the pYL2-1B sequence indicate residues identical to at least four of the five sequences; the underlined letters represent residues conserved with at least four of the five sequences. The position that is bold and underlined represents a position that is conserved but not identical amongst the five database sequences.

Alignment of pYL2-1B with Database Sequences

	1 34	39 55	65 84
pYL2-1B	**LLCD**S**CD**TP**F**H**I**Y**R**L**S**P**PL**ER**VP**S**G**D**W**I**C**xT**C**I**V**	Y**GF**T**Q**D**T**H**D**Y**S**L**P**E**F**Q**E**	L**P**ARKL**F**I**D**Ex**E**E**M**F**W**S**LV**T
	74 107	117 133	145 164
Smcy	**LLCD**G**C**SDNY**H**I**F**C**LL**PPL**SE**VP**K**G**V**W**R**C**PK**CI**L	F**GF**E**Q**A**T**Q**E**Y**T**L**Q**S**F**G**E**	M**P**VHM**V**P**T**E**V**V**E**K**E**F**W**R**LV**S
	208 241	251 267	279 297
Smcx	**LLCD**G**CD**DNY**H**I**F**C**LL**PPL**PE**I**P**K**G**V**W**R**C**PK**C**V**M**	F**GF**E**Q**A**T**R**E**Y**T**L**Q**S**F**G**E**	M**P**VHM**V**P**T**E**L**V**E**K**E**F**W**R**LV**
	305 338	348 364	376 394
MMSMCX3-1	**LLCD**G**CD**DNY**H**I**F**C**LL**PPL**PE**I**P**K**G**V**W**R**C**PK**C**V**M**	F**GF**E**Q**A**T**R**E**Y**T**L**Q**S**F**G**E**	M**P**VHM**V**P**T**E**L**V**E**K**E**F**W**R**LV**
	340 373	383 399	411 429
XE169	**LLCD**G**CD**DNY**H**I**F**C**LL**PPL**PE**I**P**K**G**V**W**R**C**PK**C**V**M**	F**GF**E**Q**A**T**R**E**Y**T**L**Q**S**F**G**E**	M**P**VHM**V**P**T**E**L**V**E**K**E**F**W**R**LV**
	309 341	352 368	380 399
RPB-2	**LLCD**G**CD**DS**Y**H**T**F**C**LI**PPL**PD**VP**K**GD**W**R**C**PK**C**V**	F**GF**E**Q**AVRE**Y**T**L**Q**S**F**G**E**	M**P**VHM**V**P**T**E**L**V**E**K**E**F**W**R**LV**S

REFERENCES

Adhya, S. and Gottesman, M. (1978). Control of transcription termination. Annu. Rev. Biochem. **47**, 967-996.

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.

Agulnik, A.I., Mitchell, M.J., Lerner, J.L., Woods, D.R., and Bishop, C.E. (1994a). A mouse Y chromosome gene encoded by a region essential for spermatogenesis and expression of male-specific minor histocompatibility antigens. Human Molecular Genetics **3**, 873-878.

Agulnik, A.I., Mitchell, M.J., Mattei, M.-G., Borsani, G., Avner, P.A., Lerner, J.L., and Bishop, C.E. (1994b). A novel X gene with a widely transcribed Y-linked homologue escapes X-inactivation in mouse and human. Human Molecular Genetics **3**, 879-884.

Ahn, B.-Y., Gershon, P.D., Jones, E.V., and Moss, B. (1990). Identification of rpo30, a vaccinia virus RNA polymerase gene with structural similarity to a eukaryotic transcription elongation factor. Mol. & Cell. Biol. **10**, 5433-5441.

Alani, E., Cao, L., and Kleckner, N. (1987). A method for gene disruption that allows repeated use of *URA3* selection in the construction of multiply disrupted yeast strains. Genetics **116**, 541-545.

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**, 599-610.

Allison, L.A., Wong, J.K.-C., Fitzpatrick, V.D., Moyle, M., and Ingles, C.J. (1988). The C-terminal domain of the largest subunit of RNA polymerase II of *Saccharomyces cerevisiae, Drosophila melanogaster*, and mammals: a conserved structure with an essential function. Mol. & Cell. Biol. **8**, 321-329.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. **215,** 403-410.

Amara, S.G., Evans, R.M., and Rosenfeld, M.G. (1984). Calcitonin/calcitonin gene-related peptide transcription unit: tissue-specific expression involves selective use of alternative polyadenylation sites. Mol. & Cell. Biol. **4**, 2151-2160.

Archambault, J., Drebot, M.A., Stone, J.C., and Friesen, J.D. (1992a). Isolation and phenotypic analysis of conditional-lethal, linker-insertion mutations in the gene encoding the largest subunit of RNA polymerase II in *Saccharomyces cerevisiae.* Mol. & Gen. Gen. **232**, 408-414.

Archambault, J., Lacroute, F., Ruet, A., and Friesen, J.D. (1992b). Genetic interaction between transcription elongation factor TFIIS and RNA polymerase II. Mol. & Cell. Biol. **12**, 4142-4152.

Archambault, J., Schappert, K.T., and Friesen, J.D. (1990). A suppressor of an RNA polymerase II mutation of *Saccharomyces cerevisiae* encodes a subunit common to RNA polymerase I, II, and III. Mol. & Cell. Biol. **10**, 6123-6131.

Arndt, K. and Fink, G.R. (1986). GCN4 protein, a positive transcription factor in yeast, binds general control promoters at all 5' TGACTC 3' sequences. Proc. Natl. Acad. Sci. USA **83**, 8516-8520.

Ashfield, R., Enriquez-Harris, P., and Proudfoot, N.J. (1991). Transcriptional termination between the closely linked human complement genes C2 and Factor B: common termination factor for C2 and c-*myc*? EMBO J. **10**, 4197-4207.

Ashfield, R., Patel, A.J., Bossone, S.A., Brown, H., Campbell, R.D., Marcu, K.B., and Proudfoot, N.J. (1994). MAZ-dependent termination between closely spaced human complement genes. EMBO J. **13**, 5656-5667.

Aso, T., Conaway, J.W., and Conaway, R.C. (1994). Role of core promoter structure in assembly of the RNA polymerase II preinitiation complex. A common pathway for formation of preinitiation intermediates at many TATA and TATA-less promoters. J. Biol. Chem. **269**, 26575-26583.

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.

Asselin, C., Nepveu, A., and Marcu, K.B. (1989). Molecular requirements for transcriptional initiation of the murine c-*myc* gene. Oncogene **4**, 549-558.

Auvinen, M., Paasinen, A., Andersson, L.C., and Holtta, E. (1992). Ornithine decarboxylase activity is critical for cell transformation. Nature **360**, 355-358.

Baek, K.H., Sato, K., Ito, R., and Agarwal, K. (1986). RNA polymerase II transcription terminates at a specific DNA sequence in a HeLa cell-free reaction. Proc. Natl. Acad. Sci. USA **83**, 7623-7627.

Baker, C.C. and Noe, J.S. (1989). Transcriptional termination between bovine papillomavirus type 1 (BPV-1) early and late polyadenylation sites blocks late transcription in BPV-1 transformed cells. J. Virology **63**, 3529-3534.

Bartel, P., Chien, C., Sternglanz, R., and Fields, S. (1993). Elimination of false positives that arise in using the two-hybrid system. BioTechniques **6**, 920-924.

Beaumont, C., Porcher, C., Picat, C., Nordmann, Y., and Grandchamp, B. (1989). The mouse porphobilinogen deaminase gene. J. Biol. Chem. **264**, 14829-14834.

Bender, T.P., Thompson, C.B., and Kuehl, W.M. (1987). Differential expression of c-*myb* mRNA in murine B lymphomas by a block to transcription elongation. Science **237**, 1473-1476.

Bengal, E. and Aloni, Y. (1991). Transcriptional elongation by purified RNA polymerase II is blocked at the *trans*-activation-response region of human immunodeficiency virus type 1 *in vitro*. J. Virology **65**, 4910-4918.

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. Mol. & Cell. Biol. **11**, 1195-1206.

Bentley, D.L., and Groudine, M. (1986). A block to elongation is largely responsible for decreased transcription of c-*myc* in differentiated HL60 cells. Nature **321**, 702-706.

Bentley, D.L. and Groudine, M. (1988). Sequence requirements for premature termination of transcription in the human c-*myc* gene. Cell **53**, 245-256.

Biragyn, A. and Nedospasov, S.A. Lipopolysaccharide-induced expression of TNF-alpha gene in the macrophage cell line ANA-1 is regulated at the level of transcription processivity. J. Immunology **155**, 674-683.

Birchmeier, C., Schumperli, D., Sconzo, G., and Birnstiel, M.L. (1984). 3' editing of mRNAs: sequence requirements and involvement of a 60-nucleotide RNA in maturation of histone mRNA precursors. Proc. Natl. Acad. Sci. USA **81**, 1057-1061.

Birnstiel, M.L., Busslinger, M., and Strub, K. (1985). Transcription termination and 3' processing: the end is in site! Cell **41**, 349-359.

Björklund, S., Skogman, E., and Thelander, L. (1992). An S-phase specific release from a transcriptional block regulates the expression of mouse ribonucleotide reductase R2 subunit. EMBO J. **11**, 4953-4959.

Blanchard, J.M., Piechaczyk, M., Dani, C., Chambard, J.C., Franchi, A., Pouyssegur, J., and Jeanteur, P. (1985). c-*myc* gene is transcribed at high rate in G0-arrested fibroblasts and is post-transcriptionally regulated in response to growth factors. Nature **317**, 443-445.

Boldin, M.P., Mett, I.L., and Wallach, D. (1995). A protein related to a proteasomal subunit binds to the intracellular domain of the p55 TNF receptor upstream to its 'death domain'. FEBS Letters **367**, 39-44.

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

Borukhov, S., Polyakov, A., Nikiforov, V., and Goldfarb, A. (1992). GreA protein : a transcription elongation factor from *Escherichia coli.* Proc. Natl. Acad. Sci. USA **89**, 8899-8902.Borukhov, S., Sagitov, V., and Goldfarb, A. (1993). Transcript cleavage factors from *Escherichia coli.* Cell **72**, 459-466.

Bossone, S.A., Asselin, C., Patel, A.J., and Marcu, K.B. (1992). MAZ, a zinc finger protein, binds to c-*myc* and C2 gene sequences regulating transcriptional initiation and termination. Proc. Natl. Acad. Sci. USA **89**, 7452-7456.

Boulikas, T. (1993). Nuclear localization signals (NLS). Critical Reviews in Eukaryotic Gene Expression **3**, 193-227.

Boulikas, T. (1994). Putative nuclear localization signals (NLS) in protein transcription factors. J. Cell. Biochem. **55**, 32-58.

Bradsher, J.N., Jackson, K.W., Conaway, R.C., and Conaway, J.W. (1993a). RNA polymerase II transcription factor SIII. I. Identification, purification, and properties. J. Biol. Chem. **268**, 25587-25593.

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

Braus, G., Paravicini, G., and Hütter, R. (1988). A consensus transcription termination sequence in the promoter region is necessary for efficient gene expression of the *TRP1* gene of *S. cerevisiae.* Mol. Gen. Genet. **212**, 495-504.

Broyles, S.S. and Moss, B. (1986). Homology between RNA polymerases of poxviruses, prokaryotes, and eukaryotes: nucleotide sequence and transcriptional analysis of vaccinia virus genes encoding 147-kDa and 22-kDa subunits. Proc. Natl. Acad. Sci. USA **83**, 3141-3145.

Buchman, A.R., Kimmerly, W.J., Rine, J., and Kornberg, R.D. (1988). Two DNA-binding factors recognize specific sequences at silencers, upstream activating sequences, autonomously replicating sequences, and telomeres in *Saccharomyces cerevisiae.* Mol. & Cell. Biol. **8**, 210-225.

Buchman, A.R. and Kornberg, R.D. (1990). A yeast *ARS*-binding protein activates transcription synergistically in combination with other weak activating factors. Mol. & Cell. Biol. **10**, 887-897.

Buck, V., Allen, K.E., Sorensen, T., Bybee, A., Hijmans, E.M., Voorhoeve, P.M., Bernards, R., and La Thangue, N.B. (1995). Molecular and functional characterization of E2F-5, a new member of the E2F family. Oncogene **11**, 31-38.

Bullock, W.O., Fernandez, J.M., and Short, J.M. (1987). XL1-blue: a high efficiency plasmid transforming recA *Escherichia coli* strain with beta-galactosidase selection. BioTechniques **5**, 376-379.

Buratowski, S. (1994). The basics of basal transcription by RNA polymerase II. Cell **77**, 1-3.

Caldwell, G.A. and Becker, J.M. (1994). Rapid PCR sequencing of plasmid DNA directly from colonies of *Saccharomyces cerevisiae.* Promega Notes **44**, 6-9.

Campbell, D.A., van Bree, M.P., and Boothroyd, J.C. (1984). The 5'-limit of transposition and upstream barren region of a trypanosome VSG gene: tandem 76 base-pair repeats flanking (TAA)90. Nucleic Acids Res. **12**, 2759-2774.

Carles, C., Treich, I., Bouet, F., Riva, M., and Sentenac, A. (1991). Two additional common subunits, ABC10α and ABC10β, are shared by yeast RNA polymerases. J. Biol. Chem. **266**, 24092-24096.

Cesarman, E., Dalla-Favera, R., Bentley, D., and Groudine, M. (1987). Mutations in the first exon are associated with altered transcription of c-*myc* in Burkitt lymphoma. Science **238**, 1272-1275.

Chamberlin, M.J. (1994). New models for the mechanism of transcription elongation and its regulation. Harvey Lectures **88**, 1-21.

Chang, C., Kostrub, C.F., and Burton, Z.F. (1993). RAP30/74 (transcription factor TFIIF) is required for promoter escape by RNA polymerase II. J. Biol. Chem. **268**, 20482-20489.

Chasman, D.I., Lue, N.F., Buchman, A.R., LaPointe, J.W., Lorch, Y., and Kornberg, R.D. (1990). A yeast protein that influences the chromatin structure of UAS_G and functions as a powerful auxiliary gene activator. Genes & Dev. **4**, 503-514.

Chaussivert, N., Conesa, C., Shaaban, S., and Sentenac A. (1995). Complex interactions between yeast TFIIIB and TFIIIC. J. Biol. Chem. **270**, 15353-15358.

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.

Chen, J.J., Reid, C.E., Band, V., and Androphy, E.J. (1995). Interaction of papillomavirus E6 oncoproteins with a putative calcium-binding protein. Science **269**, 529-531.

Chen, Z., Harless, M.L., Wright, D.A., and Kellems, R.E. (1990). Identification and characterization of transcriptional arrest sites in exon 1 of the human adenosine deaminase gene. Mol. & Cell. Biol. **10**, 4555-4564.

Chien, C., Bartel, P.L., Sternglanz, R., and Fields, S. (1991). The two-hybrid system: A method to identify and clone genes for proteins that interact with a protein of interest. Proc. Natl. Acad. Sci. USA **88**, 9578-9582.

Chinsky, J.M., Maa, M., Ramamurthy, V., and Kellems, R.E. (1989). Adenosine deaminase gene expression. J. Biol. Chem. **264**, 14561-14565.

Choder, M. (1993). A growth rate-limiting process in the last growth phase of the yeast life cycle involves RPB4, a subunit of RNA polymerase II. J. Bacteriology **175**, 6358-6363.3

Choder, M. and Young, R.A. (1993). A portion of RNA polymerase II molecules has a component essential for stress responses and stress survival. Mol. & Cell. Biol. **13**, 6984-6991.

Chodosh, L.A., Fire, A., Samuels, M., and Sharp, P.A. (1989). 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole inhibits transcription elongation by RNA polymerase II *in vitro*. J. Biol. Chem. **264**, 2250-2257.

Christie, K.R., Awrey, D.E., 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, P37. J. Biol. Chem. **269**, 936-943.

Citron, B., Falck-Pedersen, E., Salditt-Georgieff, M., and Darnell, J.E., Jr. (1984). Transcription termination occurs within a 1000 base pair region downstream from the poly(A) site of the mouse beta-globin (major) gene. Nucleic Acids Res. **12**, 8723-8731.

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. Proc. Natl. Acad. Sci. USA **91**, 8087-8091.

Ciprés-Palacín, G. and Kane, C.M. (1995). Alanine-scanning mutagenesis of human transcript elongation factor TFIIS. Biochemistry **34**, 15375-15380.

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. Mol. & Cell. Biol. **11**, 2576-2582.

Collart, M.A., Tourkine, N., Belin, D., Vassalli, P., Jeanteur, P., and Blanchard, J. (1991). c-*fos* gene transcription in murine macrophages is modulated by a calcium-dependent block to elongation in intron 1. Mol. & Cell. Biol. **11**, 2826-2831.

Conaway, J.W and Conaway, R.C. (1990). An RNA polymerase II transcription factor shares functional properties with *Escherichia coli* sigma 70. Science **248**, 1550-1553.

Conaway, R.C. and Conaway, J.W. (1993). General initiation factors for RNA polymerase II. Annu. Rev. Biochem. **62**, 161-190.

Conaway, J.W., Bradsher, J.N., Tan, S., and Conaway, R.C. (1993). Transcription factor SIII: a novel component of the RNA polymerase II elongation complex. Cell. and Mol. Biology Res. **39**, 323-329.

Connelly, S. and Manley, J.L. (1988). A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. Genes & Dev. **2**, 440-452.

Corden, J.L. (1990). Tails of RNA polymerase II. Trends Biochem. Sci. **15**, 383-387.

Cullen, B.R. (1993). Does HIV-1 Tat induce a change in viral initiation rights? Cell **73**, 417-420.

Cullen, B.R., Lomedico, P.T., and Ju, G. (1984). Transcriptional interference in avian retroviruses--implications for the promoter insertion model of leukaemogenesis. Nature **307**, 241-245.

Dahmus, M.E. (1994). The role of multisite phosphorylation in the regulation of RNA polymerase II activity. Progress in Nucleic Acid Research and Molecular Biology. **48**, 143-179.

Dedrick, R.L., Kane, C.M., and Chamberlin, M.J. (1987). Purified RNA polymerase II recognizes specific termination sites during transcription *in vitro*. J. Biol. Chem. **262**, 9098-9108.

Defeo-Jones, D., Huang, P.S., Jones, R.E., Haskell, K.M., Vuocolo, G.A., Hanobik, M.G., Huber, H.E., and Oliff, A. (1991). Cloning of cDNAs for cellular proteins that bind to the retinoblastoma gene product. Nature **352**, 251-254.

Del Sal, G., Manfioletti, G., and Schneider, C. (1989). The CTAB-DNA precipitation method: a common mini-scale preparation of template DNA from phagemids, phages, or plasmids suitable for sequencing. BioTechniques **7**, 514-520.

DesJardin, L.E., Lockwood, M.K., and Hauswirth, W.W. (1995). Bovine opsin gene expression exhibits a late fetal to adult regulatory switch. J. Neuroscience Res. **40**, 728-736.

Dieci, G., Hermann-Le Denmat, S., Lukhtanov, E., Thuriaux, P., Werner, M., and Sentenac, A. (1995). A universally conserved region of the largest subunit participates in the active site of RNA polymerase III. EMBO J. **14**, 3766-3776.

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. Proc. Natl. Acad. Sci. USA **91**, 8502-8506.

Dony, C., Kessel, M., and Gruss, P. (1985). Post-transcriptional control of *myc* and p53 expression during differentiation of the embryonal carcinoma cell line F9. Nature **317**, 636-639.

Drapkin, R. and Reinberg, D. (1994). The multifunctional TFIIH complex and transcriptional control. Trends Biochem. Sci. **19**, 504-508.

Drlica, K. (1987). "The nucleoid." in *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. F.C. Neidhardt et al., eds. (Washington, D.C.: American Society for Microbiology), pp. 91-103.

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. Proc. Natl. Acad. Sci. USA **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*. J. Biol. Chem. **266**, 71-75.

Eick, D. and Bornkamm, G.W. (1986). Transcriptional arrest within the first exon is a fast control mechanism in c-*myc* gene expression. Nucleic Acids Res. **14**, 8331-8346.

Eick, D., Kohlhuber, F., Wolf, D.A., and Strobl, L.J. (1994). Activation of pausing RNA polymerases by nuclear run-on experiments. Anal. Biochem. **218**, 347-351.

Elble, R. (1992). A simple and efficient procedure for transformation of yeasts. BioTechniques. **13**, 18-20.

Erkine, A.M., Adams, C.C., Gao, M., and Gross, D.S. (1995). Multiple protein-DNA interactions over the yeast HSC82 heat shock gene promoter. Nucleic Acids Res. **23**, 1822-1829.

Estruch, F. and Carlson, M. (1993). Two homologous zinc finger genes identified by multicopy suppression in a SNF1 protein kinase mutant of *Saccharomyces cerevisiae*. Mol. & Cell. Biol. **13**, 3872-3881.

Exinger, F. and Lacroute, F. (1992). 6-azauracil inhibition of GTP biosynthesis in *Saccharomyces cerevisiae*. Curr. Gen. **22**, 9-11.

Falck-Pederson, E., Logan, J., Shenk, T., and Darnell, J.E., Jr. (1985). Transcription termination within the E1A gene of adenovirus induced by insertion of the mouse B-major globin terminator element. Cell **40**, 897-905.

Fattaey, A.R., Helin, K., Dembski, M.S., Dyson, N., Harlow, E., Vuocolo, G.A., Hanobik, M.G., Haskell, K.M., Oliff, A., Defeo-Jones, D., and Jones, R.E. (1993). Characterization of the retinoblastoma binding proteins RBP1 and RBP2. Oncogene **8**, 3149-3156.

Feng, G., Lee, D.N., Wang, D., Chan, C.L. and Landick, R. (1994). GreA-induced transcript cleavage in transcription complexes containing *Escherichia coli* RNA polymerase is controlled by multiple factors, including nascent transcript location and structure. J. Biol. Chem. **269**, 22282-22294.

Fields, S. and Song, O. (1989). A novel genetic system to detect protein:protein interactions. Nature **340**, 245-246.

Fields, S. and Sternglanz, R. (1994). The two-hybrid system: an assay for protein:protein interactions. Trends in Genetics **10**, 286-292.

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

Frayne, E.G. and Kellems, R.E. (1986). Structural features of the murine dihydrofolate reductase transcription termination region: identification of a conserved DNA sequence element. Nucleic Acids Res. **14**, 4113-4125.

Frayne, E.G., Leys, E.J., Crouse, G.F., Hook, A.G., and Kellems, R.E. (1984). Transcription of the mouse dihydrofolate reductase gene proceeds unabated through seven polyadenylation sites and terminates near a region of repeated DNA. Mol. & Cell. Biol. **4**, 2921-2924.

Friedman, D.I., Imperiale, M.J., and Adhya, S.L. (1987). RNA 3'-end formation in the control of gene expression. Annu. Rev. Genetics **21**, 453-488.

Fort, P., Rech, J., Vie, A., Piechaczyk, M., Bonnieu, A., Jeanteur, P., and Blanchard, J. (1987). Regulation of c-*fos* gene expression in hamster fibroblasts: initiation and elongation of transcription and mRNA degradation. Nucleic Acids Res. **15**, 5657-5667.

Furter-Graves, E.M., Hall, B.D., Furter, R. (1994). Role of a small RNA pol II subunit in TATA to transcription start site spacing. Nucleic Acids Res. **22**, 4932-4936.

Galli, G., Guise, J.W., McDevitt, M.A., Tucker, P.W., and Nevins, J.R. (1987). Relative position and strengths of poly(A) sites as well as transcription termination are critical to membrane versus secreted μ-chain expression during B-cell development. Genes & Dev. **1**, 471-481.

Gariglio, P., Bellard, M., and Chambon, P. (1981). Clustering of RNA polymerase B molecules in the 5' moiety of the adult beta-globin gene of hen erythrocytes. Nucleic Acids Res. **9**, 2589-2598.

Garrett, K.P., Aso, T., Bradsher, J.N., Foundling, S.I., Lane, W.S., Conaway, R.C., and Conaway, J.W. (1995). Positive regulation of general transcription factor SIII by a tailed ubiquitin homolog. Proc. Natl. Acad. Sci. USA **92**, 7172-7176.

Garrett, K.P., Haque, D., Conaway, R.C., and Conaway, J.W. (1994a). A human cDNA encoding the small subunit of RNA polymerase II transcription factor SIII. Gene **150**, 413-414.

Garrett, K.P., Tan, S., Bradsher, J.N., Lane, W.S., Conaway, J.W., and Conaway, R.C. (1994b). Molecular cloning of an essential subunit of RNA polymerase II elongation factor SIII. Proc. Natl. Acad. Sci. USA **91**, 5237-5241.

Giardina, C., Perez-Riba, M., and Lis, J.T. (1992). Promoter melting and TFIID complexes on Drosophila genes *in vivo*. Genes & Dev. **6**, 2190-2200.

Gilmour, D.S., Buchman, A.R., and Workman, J.L. (1994). "Transcription regulation of a chromatin template." in Transcription: Mechanisms and Regulation. R.C. Conaway and J.W. Conaway, eds. (New York: Raven Press, Ltd.), pp. 515-533.

Gilmour, D.S. and Lis, J.T. (1985). *In vivo* interactions of RNA polymerase II with genes of *Drosophila melanogaster*. Mol. & Cell. Biol. **5**, 2009-2018.

Gilmour, D.S. and Lis, J.T. (1986). RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in *Drosophila melanogaster* cells. Mol. & Cell. Biol. **6**, 3984-3989.

Glenn, T.C. and Glenn, S.J. (1994). Rapid elution of DNA from agarose gels using polyester plug spin inserts (PEPSIs). Trends in Genetics **10**, 344.

Glover, J.N. and Harrison, S.C. (1995). Crystal structure of the heterodimeric bZIP transcription factor c-Fos-c-Jun bound to DNA. Nature **373**, 257-261.

Goode, B.L. and Feinstein, S.C. (1992). "Speedprep" preparation of template for double-strand DNA sequencing. BioTechniques **12**, 374-375.

Goodrich, J.A. and Tjian, R. (1994). Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. Cell **77**, 145-156.

Grant, S.G.N., Jessee, J., Bloom, F.R., and Hanahan, D. (1990). Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. Proc. Natl. Acad. Sci. USA **87**, 4645-4649.

Green, M., Maniatis, T., and Melton, D., (1983). Human β-globin pre-mRNA synthesized in vitro is accurately spliced in Xenopus oocyte nuclei. Cell **32**, 681-694.

Greenberg, M.E. and Ziff, E.B. (1984). Stimulation of 3T3 cells induces transcription of the c-*fos* proto-oncogene. Nature **311**, 433-438.

Greenblatt, J., Nodwell, J.R., and Mason, S.W. (1993). Transcriptional antitermination. Nature **364**, 401-406.

Grossman, M.E. and Tindall, D.J. (1995). The androgen receptor is transcriptionally suppressed by proteins that bind single-stranded DNA. J. Biol. Chem. **270**, 10968-10975.

Groudine, M., Peretz, M., and Weintraub, H (1981). Transcriptional regulation of hemoglobin switching in chicken embryos. Mol. & Cell. Biol. **1**, 281-288.

Gu, W, Powell, W., Mote, J., Jr., and Reines, D. (1993). Nascent RNA cleavage by arrested RNA polymerase II does not require upstream translocation of the elongation complex on DNA. J. Biol. Chem. **268**, 25604-25616.

Guo, H. and Price, D.H. (1993). Mechanism of DmS-II-mediated pause suppression by Drosophila RNA polymerase II. J. Biol. Chem. **268**, 18762-18770.

Guo, W., Wang, J., Tam, G., Yen, T.S.B., and Ou, J. (1991). Leaky transcription termination produces larger and smaller than genome size hepatitis B virus X gene transcripts. Virology **181**, 630-636.

Guo, Z., Russo, P., Yun, D., Butler, J.S., and Sherman, F. (1995). Redundant 3' end-forming signals for the yeast *CYC1* mRNA. Proc. Natl. Acad. Sci. USA **92**, 4211-4214.

Hagenbüchle, O., Wellauer, P.K., Cribbs, D.L., and Schibler, U. (1984). Termination of transcription in the mouse alpha-amylase gene Amy-2a occurs at multiple sites downstream of the polyadenylation site. Cell **38**, 737-744.

Hagler, J. and Shuman, S. (1993). Nascent RNA cleavage by purified ternary complexes of Vaccinia RNA polymerase. J. Biol. Chem. **268**, 2166-2173.

Hair, A. and Morgan, G.T. (1993). Premature termination of tubulin gene transcription in Xenopus oocytes is due to promoter-dependent disruption of elongation. Mol. & Cell. Biol. **13**, 7925-7934.

Haley, J.D. and Waterfield, M.D. (1991). Contributory effects of *de novo* transcription and premature transcript termination in the regulation of human epidermal growth factor receptor proto-oncogene RNA synthesis. J. Biol. Chem. **266**, 1746-1753.

Hall, M.N., Hereford, L., and Herskowitz, I. (1984). Targeting of E. coli beta-galactosidase to the nucleus in yeast. Cell **36**, 1057-1065.

Hanahan, D., Jessee, J., and Bloom, F.R. (1991). "Plasmid transformation of *Escherichia coli* and other bacteria." in Methods in Enzymology. J.H. Miller. ed. (San Diego, CA: Academic Press), **204**, 63-112.

Hay, N. and Aloni, Y. (1984). Attenuation in SV40 as a mechanism of transcription-termination by RNA polymerase B. Nucleic Acids Res. **12**, 1401-1414.

Hay, N., Skolnik-David, H., and Aloni, Y. (1982). Attenuation in the control of SV40 gene expression. Cell **29**, 183-193.

Hays, S.L., Firmenich, A.A., and Berg, P. (1995). Complex formation in yeast double-strand break repair: participation of Rad51, Rad52, Rad55, and Rad57 proteins. Proc. Natl. Acad. Sci. USA **92**. 6925-6929.

Hekmatpanah, D.S. and Young, R.A. (1991). Mutations in a conserved region of RNA polymerase II influence the accuracy of mRNA start site selection. Mol. & Cell. Biol. **11**, 5781-5791.

Hernandez, N. and Lucito, R. (1988). Elements required for transcription initiation of the human U2 snRNA gene coincide with elements required for snRNA 3' end formation. EMBO J. **7**, 3125-3134.

Hernandez, N. and Weiner, A.M. (1986). Formation of the 3' end of U1 snRNA requires compatible snRNA promoter elements. Cell **47**, 249-258.

Higgs, D.R., Goodbourn, S.E., Lamb, J., Clegg, J.B., Weatherall, D.J., and Proudfoot, N.J. (1983). Alpha-thalassaemia caused by a polyadenylation signal mutation. Nature **306**, 398-400.

Hill, D.E., Hope, I.A., Macke, J.P., and Struhl, K. (1986). Saturation mutagenesis of the yeast *his3* regulatory site: requirements for transcriptional induction and for binding by *GCN4* activator protein. Science **234**, 451-457.

Hirai, H., Sekimizu, K., Horikoshi, M., Nakanishi, Y., and Natori, S. (1988). Stimulation of transcription from accurate initiation sites by purified S-II. FEBS Letters **238**, 119-122.

Hirashima, S., Hirai, H., Nakanishi, Y., and Natori, S. (1988). Molecular cloning and characterization of cDNA for eukaryotic transcription factor S-II. J. Biol. Chem. **263**, 3858-3863.

Hirschhorn, R. (1993). Overview of biochemical abnormalities and molecular genetics of adenosine deaminase deficiency. Pediatric Research **33**, S35-41.

Hofer, E. and Darnell, J.E., Jr. (1981). The primary transcription unit of the mouse beta-major globin gene. Cell **23**, 585-593.

Hofer, E., Hofer-Warbinek, R., and Darnell, J.E, Jr. (1982). Globin RNA transcription: a possible termination site and demonstration of transcriptional control correlated with altered chromatin structure. Cell **29**, 887-893.

Hofmann, M.A. and Brian, D.A. (1991). Sequencing PCR DNA amplified directly from a bacterial colony. BioTechniques **11**, 30-31.

Holden, W.E. and Morgan, E.A. (1984). Antitermination of transcription from an *Escherichia coli* ribosomal RNA promoter. Proc. Natl. Acad. Sci. USA **81**, 6789-6793.

Hope, I.A. and Struhl, K. (1985). GCN4 protein, synthesized *in vitro*, binds *HIS3* regulatory sequences: implications for general control of amino acid biosynthetic genes in yeast. Cell **43**, 177-188.

Horikoshi, M., Sekimizu, K., and Natori, S. (1984). Analysis of the stimulatory factor of RNA polymerase II in the initiation and elongation complex. J. Biol. Chem. **259**, 608-611.

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. J. Biol. Chem. **260**, 5739-5744.

Horikoshi, M., Nakanishi, Y., Hirashima, S., Ohtsuki, M., Kobayashi, T., and Natori, S. (1987). Phosphorylation of S-II, a eukaryotic transcription factor, by casein kinase II. Chemical & Pharmaceutical Bulletin **35**, 4181-4187.

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. J. Biol. Chem. **265**, 11854-11857.

Hsu, S.M., Xie, S.S., el-Okda, M.O., and Hsu, P.L. (1992). Correlation of c-*fos*/c-*jun* expression with histiocytic differentiation in Hodgkin's Reed-Sternberg cells. Examination in HDLM-1 subclones with spontaneous differentiation. Amer. J. Pathology **140**, 155-165.

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

Hull, M.W., McKune, K., and Woychik, N.A. (1995). RNA polymerase II subunit RPB9 is required for accurate start site selection. Genes & Dev. **9**, 481-490.

Humphrey, T., Birse, C.E., and Proudfoot, N.J. (1994). RNA 3' end signals of the *S. pombe* ura4 gene comprise a site determining and efficiency element. EMBO J. **13**, 2441-2451.

Igarashi, K., Fujita, N., and Ishihama, A. (1990). Sequence analysis of two temperature-sensitive mutations in the alpha subunit gene (rpoA) of *Escherichia coli* RNA polymerase. Nucleic Acids Res. **18**, 5945-5948.

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

Izban, M.G. and Luse, D.S. (1993a). SII-facilitated transcript cleavage in RNA polymerase II complexes stalled early after initiation occurs in primarily dinucleotide increments. J. Biol. Chem. **268**, 12864-12873.

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

Jeon, C., Yoon, H., 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. Proc. Natl. Acad. Sci. USA **91**, 9106-9110.

Johnson, M.R., Norman, C., Reeve, M.A., Scully, J., and Proudfoot, N.J. (1986). Tripartite sequences within and 3' to the sea urchin H2A histone gene display properties associated with a transcriptional termination process. Mol. & Cell. Biol. **6**, 4008-4018.

Jones, E.W. and Fink, G.R. (1982). Regulation of amino acid biosynthesis and nucleotide biosynthesis in yeast. In The Molecular Biology of the Yeast *Saccharomyces*: metabolism and gene expression. J.N. Strathern, E.W. Jones, and J.R. Broach, eds. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory), pp. 181-299.

Jones, K.A. and Peterlin, B.M. (1994). Control of RNA initiation and elongation at the HIV-1 promoter. Annu. Rev. Biochem. **63**, 717-743.

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. J. Biol. Chem. **257**, 5286-5295.

Kanai, A., Kuzuhara, T., Sekimizu, K., and Natori, S. (1991). Heterogeneity and tissue-specific expression of eukaryotic transcription factor S-II-related protein mRNA. J. Biochem. **109**, 674-677.

Kane, C.M. (1994). "Transcript elongation and gene regulation in eukaryotes." in Transcription: Mechanisms and Regulation. R.C. Conaway and J.W. Conaway, eds. (New York: Raven Press, Ltd.), pp. 279-296.

Kao, S., Calman, A.F., Luciw, P.A., and Peterlin, B.M. (1987). Anti-termination of transcription within the long terminal repeat of HIV-1 by *tat* gene product. Nature **330**, 489-493

Kash, S.F., Innis, J.W., Jackson, A.U., and Kellems, R.E. (1993). Functional analysis of a stable transcription arrest site in the first intron of the murine adenosine deaminase gene. Mol. & Cell. Biol. **13**, 2718-2729.

Kash, S.F. and Kellems, R.E. (1994). Control of transcription arrest in intron 1 of the murine adenosine deaminase gene. Mol. & Cell. Biol. **14**, 6198-6207.

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

Kephart, D.D., Marshall, N.F., and Price, D.H. (1992). Stability of Drosophila RNA polymerase II elongation complexes *in vitro*. Mol. & Cell. Biol. **12**, 2067-2077.

Kerppola, T.K. and Kane, C.M. (1988). Intrinsic sites of transcription termination and pausing in the c-*myc* gene. Mol. & Cell. Biol. **8**, 4389-4394.

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

Kerppola, T.K. and Kane, C.M. (1991). RNA polymerase: regulation of transcript elongation and termination. FASEB J. **5**, 2833-2842.

Kessler, M., Ben-Asher, E., and Aloni, Y. (1989). Elements modulating the block of transcription elongation at the adenovirus 2 attenuation site. J. Biol. Chem. **264**, 9785-9790.

Khazak, V., Sadhale, P.P., Woychik, N.A., Brent, R., and Golemis, E.A. (1995). Human RNA polymerase II subunit hsRPB7 functions in yeast and influences stress survival and cell morphology. Mol. Biol. Cell **6**, 759-775.

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.

Kim, Y.J., Bjorklund, S., Li, Y., Sayre, M., and Kornberg, R. (1994a). A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. Cell **77**, 599-608.

Kim, Y.W., Otterson, G.A., Kratzke, R.A., Coxon, A.B., and Kaye, F.J. (1994b). Differential specificity for binding of retinoblastoma binding protein 2 to RB, p107, and TATA-binding protein. Mol. & Cell. Biol. **14**, 7256-7264.

Klopotowski, T. and Wiater, A. (1965). Synergism of aminotriazole and phosphate on the inhibition of yeast imidazole glycerol phosphate dehydratase. Arch. Biochem. Biophys. **112**, 562-566.

Koeberl, D.D., Halbert, C.L., Krumm, A., and Miller, A.D. (1995). Sequences within the coding regions of clotting factor VIII and CFTR block transcriptional elongation. Human Gene Therapy **6**, 469-479.

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

Koleske, A.J. and Young, R. (1995). The RNA polymerase II holoenzyme and its implications for gene regulation. Trends Biochem. Sci. **20**, 113-116.

Kolodziej, P.A., Woychik, N.A., Liao, S.-M., and Young, R.A. (1990). RNA polymerase II subunit composition, stoichiometry and phosphorylation. Mol. & Cell. Biol. **10**, 1915-1920.

Kolodziej, P.A. and Young, R.A. (1991). Mutations in the three largest subunits of yeast RNA polymerase II that affect enzyme assembly. Mol. & Cell. Biol. **11**, 4669-4678.

Kooter, J.M., De Lange, T., and Borst, P. (1984). Discontinuous synthesis of mRNA in trypanosomes. EMBO J. **3**, 2387-2392.

Kranz, J.E. and Holm, C. (1990). Cloning by function: an alternative approach for identifying yeast homologs of genes from other organisms. Proc. Natl. Acad. Sci. USA **87**, 6629-2233.

Krumm, A. and Groudine, M. (1995). Tumor suppression and transcription elongation: the dire consequences of changing partners. Science **269**, 1400-1401.

Krumm, A., Hickey, L.B., and Groudine, M. (1995). Promoter-proximal pausing of RNA polymerase II defines a general rate-limiting step after transcription initiation. Genes & Dev. **9**, 559-572.

Krumm, A., Meulia, T., Brunvand, M., and Groudine, M. (1992). The block to transcriptional elongation within the human c-*myc* gene is determined in the promoter-proximal region. Genes & Dev. **6**, 2201-2213.

Krumm, A., Meulia, T., and Groudine, M. (1993). Common mechanisms for the control of eukaryotic transcriptional elongation. Bioessays **15**, 659-665.

Krummel, B. and Chamberlin, M.J. (1992a). Structural analysis of ternary complexes of *Escherichia coli* RNA polymerase: Individual complexes halted along different transcription units have distinct and unexpected biochemical properties. J. Mol. Biol. **225**, 221-237.

Krummel, B. and Chamberlin, M.J. (1992b). Structural analysis of ternary complexes of *Escherichia coli* RNA polymerase: Deoxyribonuclease I footprinting of defined complexes. J. Mol. Biol. **225**, 239-250.

Krystal, G., Birrer, M., Nau, M., Sauseville, E., Thompson, C., Minna, J., and Battey, J. (1988). Multiple mechanisms for transcriptional regulation of the *myc* gene family in small-cell lung cancer. Mol. & Cell. Biol. **8**, 3373-3381.

Kuhn, A., Bartsch, I., and Grummt, I. (1990). Specific interaction of the murine transcription termination factor TTF I with class-I RNA polymerases. 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 & Dev. **3**, 224-231.

Kurschner, C. and Morgan, J.I. (1995). The cellular prion protein (PrP) selectively binds to Bcl-2 in the yeast two-hybrid system. Brain Research: Molecular Brain Research **30**, 165-8.

Laybourn, P.J. and Dahmus, M.E. (1990). Phosphorylation of RNA polymerase IIA occurs subsequent to interaction with the promoter and before the initiation of transcription. J. Biol. Chem. **265**, 13165-13173.

Landick, R. and Yanofsky, C. (1987). "Transcription attenuation." in *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. F.C. Neidhardt et al., eds. (Washington, D.C.: American Society for Microbiology), pp. 1276-1301.

Lee, D.N. and Landick, R. (1992). Structure of RNA and DNA chains in paused transcription complexes containing *Escherichia coli* RNA polymerase. J. Mol. Biol. **228**, 759-777.

Lee, G. and Gilman, M. (1994). Dual modes of control of c-*fos* mRNA induction by intracellular calcium in T cells. Mol. & Cell. Biol. **14**, 4579-4587.

Lee, H., Kraus, K.W., Wolfner, M.F., and Lis, J.T. (1992). DNA sequence requirements for generating paused polymerase at the start of *hsp70*. Genes & Dev. **6**, 284-295.

LeMeur, M.A., Galliot, B., and Gerlinger, P (1984). Termination of the ovalbumin gene transcription. EMBO J. **3**, 2779-2786.

Li, Y. and Kornberg, R.D. (1994). Interplay of positive and negative effectors in function of the C-terminal repeat domain of RNA polymerase II. Proc. Natl. Acad. Sci. USA **91**, 2362-2366.

Li, S.C., Squires, C.L., and Squires, C. (1984). Antitermination of *E. coli* rRNA transcription is caused by a control region segment containing lambda *nut*-like sequences. Cell **38**, 851-860.

Linial, M., Gunderson, N., and Groudine, M. (1985). Enhanced transcription of c-*myc* in bursal lymphoma cells requires continuous protein synthesis. Science **230**, 1126-1132.

Linn, S.C. and Luse, D.S. (1991). RNA polymerase II elongation complexes paused after the synthesis of 15- or 35-base transcripts have different structures. Mol. & Cell. Biol. **11**, 1508-1522.

Liu, F., Ventura, F., Doody, J., and Massague, J. (1995). Human type II receptor for bone morphogenic proteins (BMPs): extension of the two-kinase receptor model to the BMPs. Mol. & Cell. Biol. **15**, 3479-3486.

Lois, R., Freeman, L., Villeponteau, B., and Martinson, H.G. (1990). Active β-globin gene transcription occurs in methylated, DNase I-resistant chromatin of nonerythroid chicken cells. Mol. & Cell. Biol. **10**, 16-27.

Lu, X., Welsh, T., Peterlin, M.B. (1993). The human immunodeficiency virus type 1 long terminal repeat specifies two different complexes, only one of which is regulated by tat. J. Virology **67**, 1752-1760.

Lüscher, B. and Eisenman, R.N. (1990a). New light on *myc* and *myb*. Part I. *Myc*. Genes & Dev. **4**, 2025-2035.

Lüscher, B. and Eisenman, R.N. (1990b). New light on *myc* and *myb*. Part II. *Myb*. Genes & Dev. **4**, 2235-2241.

Ma, J. and Ptashne, M. (1987). Deletion analysis of GAL4 defines two transcriptional activating segments. Cell **48**, 847-853.

Maderious, A. and Chen-Kiang, S. (1984). Pausing and premature termination of human RNA polymerase II during transcription of adenovirus *in vivo* and *in vitro*. Proc. Natl. Acad. Sci. USA **81**, 5931-5935.

Maldonado, E. and Reinberg, D. (1995). News on initiation and elongation of transcription by RNA polymerase II. Curr. Opin. Cell Biol. **7**, 352-361.

Marcu, K.B., Bossone, S.A., and Patel, A.J. (1992). *myc* function and regulation. Annu. Rev. Biochem. **61**, 809-860.

Marshall, N.F. and Price, D.H. (1992) Control of formation of two distinct classes of RNA polymerase II elongation complexes. Mol. & Cell. Biol. **12**, 2078-2090.

Marshall, N.F. and Price D.H. (1995). Purification of P-TEFb, a transcription factor required for the transition into productive elongation. J. Biol. Chem. **270**, 12335-12338.

Marshall, T.K., Guo, H., and Price, D.H. (1990). Drosophila RNA polymerase II elongation factor DmS-II has homology to mouse S-II and sequence similarity to yeast *PPR2*. Nucleic Acids Res. **18**, 6293-6298.

Mather, E.L., Nelson, K.J., Haimovich, J., and Perry, R.P. (1984). Mode of regulation of immunoglobulin mu and delta chain expression varies during B lymphocyte maturation. Cell **36**, 329-338.

Maxon, M.E., Goodrich, J.A., and Tjian, R. (1994). Transcription factor IIE binds preferentially to RNA polymerase IIa and recruits TFIIH: a model for promoter clearance. Genes & Dev. **8**, 515-524.

McBroom, L.D. and Sadowski, P.D. (1994a). Contacts of the ABF1 protein of *Saccharomyces cerevisiae* with a DNA binding site at MATa. J. Biol. Chem. **269**, 16455-16460.

McBroom, L.D and Sadowski, P.D. (1994b). DNA bending by *Saccharomyces cerevisiae* ABF1 and its proteolytic fragments. J. Biol. Chem. **269**, 16461-16468.

McCann, P.P. and Pegg, A.E. (1992). Ornithine decarboxylase as an enzyme target for gene therapy. Pharmacology and Therapeutics **54**, 195-215.

Mechti, N., Piechaczyk, M., Blanchard, J., Jeanteur, P., and Lebleu, B. (1991). Sequence requirements for premature transcription arrest within the first intron of the mouse c-*fos* gene. Mol. & Cell. Biol. **11**, 2832-2841.

McKune, K., Richards, K.L., Edwards, A.M., and Kornberg, R.D. (1993). RBP7, one of two dissociable subunits of yeast RNA polymerase II, is essential for cell viability. Yeast **9**, 295-299.

McKune, K. and Woychik, N.A. (1994a). Halobacterial S9 operon contains two genes encoding proteins homologous to subunits shared by eukaryotic RNA polymerases I, II, and III. J. Bacteriology **176**, 4754-4756.

McKune, K. and Woychik, N.A. (1994b). Functional substitution of an essential yeast RNA polymerase subunit by a highly conserved mammalian counterpart. Mol. & Cell. Biol. **14**, 4155-4159.

Mendelsohn, C., Larkin, S., Mark, M., LeMeur, M., Clifford, J., Zelent, A., and Chambon, P. (1994). RAR β isoforms: distinct transcriptional control by retinoic acid and specific spatial patterns of promoter activity during mouse embryonic development. Mechanisms of Dev. **45**, 227-241.

Meulia, T., Krumm, A., and Groudine, M. (1993). Distinct properties of c-*myc* transcriptional elongation are revealed in Xenopus oocytes and mammalian cells and by template titration, 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB), and promoter mutagenesis. Mol. & Cell. Biol. **13**, 5647-5658.

Meulia, T., Krumm, A., Spencer, C., and Groudine, M. (1992). Sequences in the human c-*myc* P2 promoter affect the elongation and premature termination of transcripts initiated from the upstream P1 promoter. Mol. & Cell. Biol. **12**, 4590-4600.

Middleton, K.M. and Morgan, G.T. (1990). Premature termination of transcription can be induced on an injected α-tubulin gene in *Xenopus* oocytes. Mol. & Cell. Biol. **10**, 727-735.

Miller, H., Asselin, C., Dufort, D., Yang, J.Q., Gupta, K., Marcu, K.B., and Nepveu, A. (1989). A cis-acting element in the promoter region of the murine c-*myc* gene is necessary for transcriptional block. Mol. & Cell. Biol. **9**, 5340-5349.

Minvielle-Sebastia, L., Winsor, B., Bonneaud, N., and Lacroute, F. (1991). Mutations in the yeast *RNA14* and *RNA15* genes result in an abnormal mRNA decay rate; sequence analysis reveals an RNA-binding domain in the *RNA15* protein. Mol. & Cell. Biol. **11**, 3075-3087.

Mirkovitch, J. and Darnell, J.E., Jr. (1991). Rapid *in vivo* footprinting technique identifies proteins bound to the TTR gene in the mouse liver. Genes & Dev. **5**, 83-93.

Mirkovitch, J. and Darnell, J.E., Jr. (1992). Mapping of RNA polymerase on mammalian genes in cells and nuclei. Mol. Biol. Cell **3**, 1085-1094.

Mirkovitch, J., Decker, T., and Darnell, J.E., Jr. (1992). Interferon induction of gene transcription analyzed by *in vivo* footprinting. Mol. & Cell. Biol. **12**, 1-9.

Moberg, K.H., Logan, T.J., Tyndall, W.A., and Hall, D.J. (1992). Three distinct elements within the murine c-*myc* promoter are required for transcription. Oncogene **7**, 411-421.

Mok, M., Maderious, A., and Chen-Kiang, S. (1984). Premature termination by human RNA polymerase II occurs temporally in the adenovirus major late transcriptional unit. Mol. & Cell. Biol. **4**, 2031-2040.

Moreira, A., Wollerton, M., Monks, J., and Proudfoot, N.J. (1995). Upstream sequence elements enhance poly(A) site efficiency of the C2 complement gene and are phylogenetically conserved. EMBO J. **14**, 3809-3819.

Morgan, E. (1986). Antitermination mechanisms in rRNA operons of *Escherichia coli.* J. Bacteriology **168**, 1-5.

Mota, M. and Soares, E.V. (1994). Population dynamics of flocculating yeasts. FEMS Microbiology Reviews **14**, 45-52.

Myers, A.M., Crivellone, M.D., Koerner, T.J. and Tzagoloff, A. (1987). Characterization of the yeast *HEM2* gene and transcriptional regulation of COX5 and COR1 by heme. J. Biol. Chem. **262**, 16822-16829

Mytelka, D.S. (1995). Regulation of motility in *Escherichia coli.* Ph.D. Thesis, University of California, Berkeley.

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.* J. Biol. Chem. **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, and suppression of 6-azauracil sensitivity. **270**, 8991-8995.

Natori, S. (1982). Stimulatory proteins of RNA polymerase II from Ehrlich ascites tumor cells. Mol. & Cell. Biochem. **46**, 173-187.

Nepveu, A. and Marcu, K.B. (1986). Intragenic pausing and anti-sense transcription within the murine c-*myc* locus. EMBO J. **5**, 2859-2865.

Nepveu, A., Marcu, K.B., Skoultchi, A.I., and Lachman, H.M. (1987). Contributions of transcriptional and post-transcriptional mechanisms to the regulation of c-*myc* expression in mouse erythroleukemia cells. Genes & Dev. **1**, 938-945.

Neuman de Vegvar, H.E., Lund, E., and Dahlberg, J.E. (1986). 3' end formation of U1 snRNA precursors is coupled to transcription from snRNA promoters. Cell **47**, 259-266.

Nishikura, N. (1986). Sequences involved in accurate and efficient transcription of human c-*myc* genes microinjected into frog oocytes. Mol. & Cell. Biol. **6**, 4093-4098.

Nonet, M.L., Sweetser, D., and Young, R.A. (1987). Functional redundancy and structural polymorphism in the large subunit of RNA polymerase II. Cell **50**, 909-915.

Nonet, M.L. and Young, R.A. (1989). Intragenic and extragenic suppressors of mutations in the heptapeptide repeat domain of *Saccharomyces cerevisiae* RNA polymerase II. Genetics **123**, 715-724.

Nudler, E., Goldfarb, A., and Kashlev. M. (1994). Discontinuous mechanism of transcription elongation. Science **265**, 793-796.

Orkin, S.H., Cheng, T.C., Antonarakis, S.E., and Kazazian, H.H., Jr. (1985). Thalassemia due to a mutation in the cleavage-polyadenylation signal of the human beta-globin gene. EMBO J. **4**, 453-456.

Orlova, M., Newlands, J., Das, A., Goldfarb, A., and Borukhov. S. (1995). Intrinsic transcript cleavage activity of RNA polymerase. Proc. Natl. Acad. Sci. USA **92**, 4596-4600.

O'Shea, E.K., Rutkowski, R., and Kim, P.S. (1992). Mechanism of specificity in the Fos-Jun oncoprotein heterodimer. Cell **68**, 699-708.

Ozanne, B., Richards, C.S., Hendler, F., Burns, D., and Gusterson, B. (1986). Over-expression of the EGF receptor is a hallmark of squamous cell carcinomas. J. Pathology **149**, 9-14.

Parry, H.D., Tebb, G., and Mattaj, I.W. (1989). The Xenopus U2 gene PSE is a single, compact, element required for transcription initiation and 3' end formation. Nucleic Acids Res. **17**, 3633-3644.

Pegg, A.E. (1986). Recent advances in the biochemistry of polyamines in eukaryotes. Biochem. J. **234**, 249-262.

Pinto, I., Ware, D.E., and Hampsey, M. (1992). The yeast *SUA7* gene encodes a homolog of human transcription factor TFIIB and is required for normal start site selection *in vivo*. Cell **68**, 977-988.

Plet, A., Eick, D., and Blanchard, J.M. (1995). Elongation and premature termination of transcripts initiated from c-*fos* and c-*myc* promoters show dissimilar patterns. Oncogene **10**, 319-328.

Porcher, C., Picat, C., Daegelen, D., Beaumont, C., and Grandchamp, B. (1995). Functional analysis of DNase-I hypersensitive sites at the mouse porphobilinogen deaminase gene locus. J. Biol. Chem. **270**, 17368-17374.

Preker, P.J., Lingner, J., Minvielle-Sebastia, L., and Keller, W. (1995). The *FIP1* gene encodes a component of a yeast pre-mRNA polyadenylation factor that directly interacts with poly(A) polymerase. Cell **81**, 379-389.

Price, D.H., Sluder, A.E., and Greenleaf, A.L. (1987). Fractionation of transcription factors for RNA polymerase II from Drosophila Kc cell nuclear extracts. J. Biol. Chem. **262**, 3244-3255.

Price, D.H., Sluder, A.E., and Greenleaf, A.L. (1989). Dynamic interaction between a Drosophila transcription factor and RNA polymerase II. Mol. & Cell. Biol. **9**, 1465-1475.

Proudfoot, N.J. (1986). Transcriptional interference and termination between duplicated α-globin gene constructs suggests a novel mechanism for gene regulation. Nature **322**, 562-565.

Proudfoot, N.J. (1989). How RNA polymerase II terminates transcription in higher eukaryotes. Trends Biochem. Sci. **14**, 105-110.

Puhler, G., Leffers, H., Gropp, F., Palm, P., Klenk, H., Lottspeich, F., Garrett, R.A., and Zillig, W. (1989). Archaebacterial DNA-dependent RNA polymerases testify to the evolution of the eukaryotic nuclear genome. Proc. Natl. Acad. Sci. USA **86**, 4569-4573.

Pyrc, J.J., Moberg, K.H., and Hall, D.J. (1992). Isolation of a novel cDNA encoding a zinc-finger protein that binds to two sites within the c-*myc* promoter. Biochemistry **31**, 4102-4110.

Qian, X., Gozani, S.N., Yoon, H., Jeon, C.J., Agarwal, K., Weiss, M.A. (1993a). Novel zinc finger motif in the basal transcriptional 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., Weiss, M.A. (1993b). Structure of a new nucleic-acid-binding motif in eukaryotic transcriptional elongation factor TFIIS. Nature **365**, 277-279. Published correction: **376**, 279.

Rappaport, J., Reinberg, D., Zandomeni, R., and Weinmann, R. (1987). Purification and functional characterization of transcription factor SII from calf thymus. J. Biol. Chem. **262**, 5227-5232.

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

Rasmussen, E.B. and Lis, J.T. (1995). Short transcripts of the ternary complex provide insight into RNA polymerase II elongational pausing. J. Mol. Biol. **252**, 522-535.

Ratnasabapathy, R., Sheldon, M., Johal, L., and Hernandez, N. (1990). The HIV-1 long terminal repeat contains an unusual element that induces the synthesis of short RNAs from various mRNA and snRNA promoters. Genes & Dev. **4**, 2061-2074.

Reinberg, D. and Roeder, R.G. (1987). Factors involved in specific transcription by mammalian RNA polymerase II. J. Biol. Chem. **262**, 3331-3337.

Reines, D. (1992). Elongation factor-dependent transcript shortening by template-engaged RNA polymerase II. J. Biol. Chem. **267**, 3795-3800.

Reines, D.R. (1994). "Nascent RNA cleavage by elongation complexes." in Transcription: Mechanisms and Regulation. R.C. Conaway and J.W. Conaway, eds. (New York: Raven Press, Ltd.), pp. 263-278.

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

Reines, D., Ghanouni, P., Li, Q., and Mote, J., Jr. (1992). The RNA polymerase II elongation complex: factor-dependent transcription involves nascent RNA cleavage. J. Biol. Chem. **267**, 15516-15522.

Reines, D. and Mote, J., Jr. (1993). Elongation factor SII-dependent transcription by RNA polymerase II through a sequence specific DNA-binding protein. Proc. Natl. Acad. Sci. USA **90**, 1917-1920.

Reines, D., Wells, D., Chamberlin, M.J., and Kane, C.M. (1987). Identification of intrinsic termination sites *in vitro* for RNA polymerase II within eukaryotic gene sequences. J. Mol. Biol. **196**, 299-312.

Remacle, J.E. and Holmberg, S. (1992). A REB1-binding site is required for GCN4-independent *ILV1* basal level transcription and can be functionally replaced by an ABF1-binding site. Mol. & Cell. Biol. **12**, 5516-5526.

Resnekov, O. and Aloni, Y. (1989). RNA polymerase II is capable of pausing and prematurely terminating transcription at a precise location both *in vivo* and *in vitro*. Proc. Natl. Acad. Sci. USA **86**, 12-16.

Resnekov, O., Kessler, M., and Aloni, Y. (1989). RNA secondary structure is an integral part of the *in vitro* mechanism of attenuation in simian virus 40. J. Biol. Chem. **264**, 9953-9959.

Rice, G.A., Chamberlin, M.J., and Kane, C.M. (1993). Contacts between mammalian RNA polymerase II and the template DNA in a ternary elongation complex. Nucleic Acids Res. **21**, 113-118.

Ring, B.Z. and Roberts, R.W. (1994). Function of a nontranscribed DNA strand site in transcription elongation. Cell **78**, 317-324.

Roberts, J.W. (1993). RNA and protein elements of *E. coli* and λ transcription antitermination complexes. Cell **72**, 653-655.

Roberts, S. and Bentley, D.L. (1992). Distinct modes of transcription read through or terminate at the c-*myc* attenuator. EMBO J. **11**, 1085-1093.

Roberts, S., Purton, T., and Bentley, D.L. (1992). A protein-binding site in the c-*myc* promoter functions as a terminator of RNA polymerase II transcription. Genes & Dev. **6**, 1562-1574.

Rohrbaugh, M.L., Johnson, J.E., 3d, James, M.D., and Hardison, R.C. (1985). Transcription unit of the rabbit beta 1 globin gene. Mol. & Cell. Biol. **5**, 147-160.

Ross, I.L., Dunn, T.L., Yue, X., Roy, S., Barnett, C.J.K., and Hume, D.A. (1994). Comparison of the expression and function of the transcription factor PU.1 (SPI-1 proto-oncogene) between murine macrophages and B lymphocytes. Oncogene **9**, 121-132.

Rougvie, A.E. and Lis, J.T. (1988). The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of *D. melanogaster* is transcriptionally engaged. Cell **54**, 795-804.

Rougvie, A.E. and Lis, J.T. (1990). Postinitiation transcriptional control in *Drosophila melanogaster*. Mol. & Cell. Biol. **10**, 6041-6045.

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. Proc. Natl. Acad. Sci. USA **91**, 8057-8061.

Ruet, A., Riva, M., and Sentenac, A. (1990). P37 is homologous to the murine transcription elongation factor S-II. Yeast **6**, p. 05-36A.

Sasse-Dwight, S. and Gralla, J.D. (1988). Probing the *E. coli* ginALG upstream activation mechanism *in vivo*. Proc. Natl. Acad. Sci. USA **85**, 8934-8938.

Sasse-Dwight, S. and Gralla, J.D. (1989). $KMnO_4$ as a probe for lac promoter DNA melting and mechanism *in vivo*. J. Biol. Chem. **264**, 8074-8081.

Sasse-Dwight, S. and Gralla, J.D. (1990). Role of eucaryotic-type functional domains found in the prokaryotic enhancer receptor factor σ54. Cell **62**, 945-954.

Sato, K., Ito, R., Baek, K.H., and Agarwal, K. (1986). A specific DNA sequence controls termination of transcription in the gastrin gene. Mol. & Cell. Biol. **6**, 1032-1043.

Sawadogo, M., Huet, J., and Fromageot, P. (1980). Similar binding site for P37 factor on yeast RNA polymerases A and B. Biochem. Biophys. Res. Commun. **96**, 258-264.

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

Sawadogo, M. and Sentenac, A. (1990). RNA polymerase B(II) and general transcription factors. Annu. Rev. Biochem. **59**, 711-754.

Sawadogo, M., Sentenac, A., and Fromageot, P. (1979). Interaction of a new polypeptide with yeast RNA polymerase B. J. Biol. Chem. **255**, 12-15.

Scafe, C., Chao, D., Lopes, J., Hirsch, J.P., Henry, S., and Young, R.A. (1990). RNA polymerase II C-terminal repeat influences response to transcriptional enhancer signals. Nature **347**, 491-494.

Schiestl, R.H. and Gietz, D.R. (1989). High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. Curr. Genet. **16**, 339-346.

Schlaepfer, I.R., Mattoon, J.R., and Bajszar, G. (1994). The sequence and potential regulatory elements of the *HEM2* promoter of *Saccharomyces cerevisiae.* Yeast **10**, 227-229.

Schnier, J., Schwelberger, H.G., Smit-McBride, Z., Kang, H.A., and Hershey, J.W.B. (1991). Translation initiation factor 5A and its hypusine modification are essential for cell viability in the yeast *S. cerevisiae.* Mol. & Cell. Biol. **11**, 3105-3114.

Scott, D.M., Ehrmann, I.E., Ellis, P.S., Bishop, C.E., Agulnik, A.I., Simpson, E., and Mitchell, M.J. (1995). Identification of a mouse male-specific transplantation antigen, H-Y. Nature **376**, 695-698.

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.

Sekimizu, K., Mizuno, D., and Natori, S. (1979a). Localization of a factor stimulating RNA polymerase II in the nucleoplasm, revealed by immunofluorescence. Experimental Cell Research **124**, 63-72.

Sekimizu, K., Nakanishi, Y., Mizuno, D., and Natori, S. (1979b). Purification and preparation of antibody to RNA polymerase II stimulatory factors from Ehrlich ascites tumor cells. Biochemistry **18**, 1582-1588.

Sekimizu, K., Kubo, Y., Segawa, K., and Natori, S. (1981). Difference in phosphorylation of two factors stimulating RNA polymerase II of Ehrlich ascites cells. Biochemistry **20**, 2286-2292.

Sekimizu, K., Horikoshi, M., and Natori, S. (1984). Enhancement of the frequency of initiation by a stimulatory protein of RNA polymerase II. J. Biochem. **96**, 1859-1865.

Selleck, S.B. and Majors, J. (1987). Photofootprinting *in vivo* detects transcription-dependent changes in yeast TATA boxes. Nature **325**, 173-177.

Sentenac, A. (1985). Eukaryotic RNA polymerases. CRC Crit. Rev. Biochem. **18**, 31-90.

Shaaban, S.A., Krupp, B.M., and Hall, B.D. (1995). Termination-altering mutations in the second-largest subunit of yeast RNA polymerase III. Mol. & Cell. Biol. **15**, 1467-1478.

Sharrock, R.A., Gourse, R.L., and Nomura, M. (1985). Defective antitermination of rRNA transcription and derepression of rRNA and tRNA synthesis in the *nusB5* mutant of *Escherichia coli.* Proc. Natl. Acad. Sci. USA **82**, 5275-5279.

Sheffery, M., Marks, P.A., and Rifkind, R.A. (1984). Gene expression in murine erythroleukemia cells: transcriptional control and chromatin structure of the α-globin gene. J. Mol. Biol. **172**, 417-436.

Sheldon, M., Ratnasabapathy, R., and Hernandez, N. (1993). Characterization of the inducer of short transcript, a human immunodeficiency virus type 1 transcriptional element that activates the synthesis of short RNAs. Mol. & Cell. Biol. **13**, 1251-1263.

Sheldon, M. and Reinberg, D. (1995). Transcriptional Activation. Tuning-up transcription. Current Biol. **5**, 43-46.

Shiekhattar, R., Mermelstein, F., Fisher, R.P., Drapkin, R., Dynlacht, B., Wessling, H.C., Morgan, D.O., and Reinberg, D. (1995). Cdk-activating kinase complex is a component of human transcription factor TFIIH. Nature **374**, 283-287.

Shor, J., Ben-Asher, E., and Aloni, Y. (1995). Transcription elongation of the murine ornithine decarboxylase (ODC) gene is regulated *in vitro* at two downstream elements by different attenuation mechanisms. Oncogene **10**, 1587-1596.

Shpakovski, G.V., Acker, J., Wintzerith, M., Lacroix, J., Thuriaux, P., and Vigneron, M. (1995). Four subunits that are shared by the three classes of RNA polymerase are functionally interchangeable between *Homo sapiens* and *Saccharomyces cerevisiae.* Mol. & Cell. Biol. **15**, 4702-4710.

Sikorski, R.S. and Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae.* Genetics **122**, 19-27.

SivaRaman, L., Reines, D., and Kane, C.M. (1990). Purified elongation factor SII is sufficient to promote read-through by purified RNA polymerase II at specific termination sites in the human histone H3.3 gene. J. Biol. Chem. **265**, 14554-14560.

Skarnes, W.C., Tessier, D.C., and Acheson, N.H. (1988). RNA polymerases stall and/or prematurely terminate nearby both early and late promoters on polyomavirus DNA. J. Mol. Biol. **203**, 153-71.

Sluder, A.E., Greenleaf, A.L., and Price, D.H. (1989). Properties of Drosophila RNA polymerase II elongation factor. J. Biol. Chem. **264**, 8963-8969.

Smit, G., Straver, M.H., Lugtenberg, B.J.J., and Kijne, J.W. (1992). Flocculence of *Saccharomyces cerevisiae* cells is induced by nutrient limitation, with cell surface hydrophobicity as a major determinant. Applied and Environmental Microbiology **58**, 3709-3714.

Spencer, C.A and Kilvert, M.A. (1993). Transcription elongation in the human c-*myc* gene is governed by overall transcription initiation levels in Xenopus oocytes. Mol. & Cell. Biol. **13**, 1296-1305.

Spencer, C.A., LeStrange, R.C., Novak, U., Hayward, W.S., and Groudine, M. (1990). The block to transcription elongation is promoter dependent in normal and Burkitt's lymphoma c-*myc* alleles. Genes & Dev. **4**, 75-88.

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

Squires, C.L., Greenblatt, J., Li, J., Condon, C., and Squires, C.L. (1993). Ribosomal RNA antitermination *in vitro*: requirement for Nus factors and one or more unidentified cellular components. Proc. Natl. Acad. Sci. USA **90**, 970-974.

Stewart, V. and Yanofsky, C. (1985). Evidence for transcription antitermination control of tryptophanase gene expression in *Escherichia coli* K-12. J. Bacteriology **164**, 731-740.

Stratford, M. (1992a). Lectin-mediated aggregation of yeasts - yeast flocculation. Biotechnology and Genetic Engineering Reviews **10**, 283-341.

Stratford M. (1992b). Yeast flocculation: a new perspective. Advances in Microbial Physiology. **33**, 2-71.

Strobl, L.J. and Eick, D. (1992). Hold back of RNA polymerase II at the transcription start site mediates down-regulation of c-*myc in vivo*. EMBO J. **11**, 3307-3314.

Strobl, L.J., Kohlhuber, F., Mautner, J., Polack, A., and Eick, D. (1993). Absence of a paused transcription complex from the c-*myc* P2 promoter of the translocation chromosome in Burkitt's lymphoma cells: implication for the c-*myc* P1/P2 promoter shift. Oncogene **8**, 1437-1447.

Struhl, K. (1989). Molecular mechanisms of transcriptional regulation in yeast. Annu. Rev. Biochem. **58**, 1051-1077.

Struhl, K. and Davis, R.W. (1977). Production of a functional eukaryotic enzyme in *Escherichia coli*: cloning and expression of the yeast structural gene for imidazole-glycerolphosphate dehydratase (his3). Proc. Natl. Acad. Sci. USA **74**, 5255-5259.

Su, L.K., Burrell, M., Hill, D.E., Gyuris, J., Brent, R., Wiltshire, R. Trent, J., Vogelstein, B., and Kinzler, K.W. (1995). APC binds to the novel protein EB1. Cancer Research **55**, 2972-2977.

Surratt, C.K., Milan, S.C., and Chamberlin, M.J. (1991). Spontaneous cleavage of RNA in ternary complexes of Escherichia coli RNA polymerase and its significance for the mechanism of transcription. Proc. Natl. Acad. Sci. USA **88**, 7983-7987.

Sweetser, D., Nonet, M.L., and Young, R.A. (1987). Prokaryotic and eukaryotic RNA polymerases have homologous core subunits. Proc. Natl. Acad. Sci. USA **84**, 1192-1196.

Takagi, Y., Conaway, J.W., and Conaway, R.C. (1995). A novel activity associated with RNA polymerase II elongation factor SIII. J. Biol. Chem. **270**, 24300-24305.

Takeuchi, T., Yamazaki, Y., Katoh-Fukui, Y., Tsuchiya, R., Kondo, S., Motoyama, J., and Higashinakagawa, T. (1995). Gene trap capture of a novel mouse gene, *jumonji*, required for neural tube formation.

Tan, S., Conaway, R.C., and Conaway, J.W. (1995). Dissection of transcription factor TFIIF functional domains required for initiation and elongation. Proc. Natl. Acad. Sci. USA **92**, 6042-6046.

Thompson, N.E., Aronson, D.B., and Burgess, R.R. (1990). Purification of eukaryotic RNA polymerase II by immunoaffinity chromatography: Elution of active enzyme with protein stabilizing agents from a polyol-responsive monoclonal antibody. J. Biol. Chem. **265**, 7069-7077.

Thuriaux, P. and Sentenac, A. (1992). "Yeast nuclear RNA polymerases" in The Molecular and Cellular Biology of the Yeast *Saccharomyces*. Volume II.: Gene Expression. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 1-48.

Tjian, R. (1995). Molecular mechanisms that control genes. Scientific American. **272**, 54-61.

Tjian, R. and Maniatis, T. (1994). Transcriptional activation: a complex puzzle with a few easy pieces. Cell **77**, 5-8.

Treich I, Cairns, B,R., de los Santos, T., Brewster, E., and Carlson, M. (1995). SNF11, a new component of the yeast SNF-SWI complex that interacts with a conserved region of SNF2. Mol. & Cell. Biol. **15**, 4240-4248.

Tong, X., Wang, F., Thut, C.J., and Kieff, E. (1995a). The Epstein-Barr virus nuclear protein 2 acidic domain can interact with TFIIB, TAF40, and RPA70 but not with TATA-binding protein. J. Virology **69**, 585-588.

Tong, X., Drapkin, R., Reinberg, D., and Kieff, E. (1995b). The 62- and 80-kDa subunits of transcription factor IIH mediate the interaction with Epstein-Barr virus nuclear protein 2. Proc. Natl. Acad. Sci. USA **92**, 3259-3263.

Toth, C.R., Hostutler, R.F., Baldwin, A.S., and Bender, T.P. (1995). Members of the nuclear factor KB family transactivate the murine c-*myb* gene. J. Biol. Chem. **270**, 7661-7671.

Turnbough, C.L. and Landick, R (1992). in Transcriptional Attenuation. S.L. McKnight and K.R. Yamamoto, eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 407-446.

Ueno, K., Sekimizu, K., Mizuno, D., and Natori, S. (1979). Antibody against a stimulatory factor of RNA polymerase II inhibits nuclear RNA synthesis. Nature **277**, 145-146.

Usheva, A., Maldonado, E., Goldring, A., Lu, H., Houbavi, C., Reinberg, D., and Aloni, Y. (1992). Specific interaction between the nonphosphorylated form of RNA polymerase II and the TATA-binding protein. Cell **69**, 871-881.

Valay, J., Simon, M., Dubois, M., Bensaude, O., Facca, C., and Faye, G. (1995). The *KIN28* gene is required both for RNA polymerase II mediated transcription and phosphorylation of the Rpb1p CTD. J. Mol. Biol. **249**, 535-544.

Vandenbergh, D.J., James-Pederson, M., and Hardison, R.C. (1991). An apparent pause site in the transcription unit of the rabbit alpha-globin gene. J. Mol. Biol. **220**, 255-270.

van Holde K. (1993). The omnipotent nucleosome. Nature **362**, 111-112.

Vazquez de Aldana, C.R., Marton, M.J., and Hinnebusch, A.G. (1995). GCN20, a novel ATP binding cassette protein, and GCN1 reside in a complex that mediates activation of the eIF-2 alpha kinase GCN2 in amino acid-starved cells. EMBO J. **14,** 3184-3199.

Villeponteau, B., Landes, G.M., Pankratz, M.J., and Martinson, H.G. (1982). The chicken beta globin gene region. Delineation of transcription units and developmental regulation of interspersed DNA repeats. J. Biol. Chem. **257**, 11015-11023.

von Hippel, P.H. and Yager, T.D. (1992). The elongation-termination decision in transcription. Science **255**, 809-812.

Wang, B.Q. and Burton, Z.F. (1995). Functional domains of human RAP74 including a masked polymerase binding domain. J. Biol. Chem. **270**, 27035-27044.

Wang, D. and Hawley, D.K. (1993). Identification of a 3' to 5' exonuclease activity associated with human RNA polymerase II. Proc. Natl. Acad. Sci. USA **90**, 843-847.

Wang, W., Meadows, L.R., den Haan, J.M.M., Sherman, N.E., Chen, Y., Blokland, E., Shabanowitz, J., Agulnik, A.I., Hendrickson, R.C., Bishop, C.E., Hunt, D.F., Goulmy, E., and Engelhard, V.H. (1995). Human H-Y: a male-specific histocompatibility antigen derived from the SMCY protein. Science **269**, 1588-1590.

Ward, A.C. (1990). Single-step purification of shuttle vectors from yeast for high frequency back-transformation into *E. coli.* Nucleic Acids Res. **18**, 5319.

Watson, R. (1988). Expression of the c-*myb* and c-*myc* genes is regulated independently in differentiating mouse erythroleukemia cells by common processes of premature transcription arrest and increased mRNA turnover. Mol. & Cell. Biol. **8**, 3938-3942.

Weaver, Z.A. (1995). Identification and regulation of the human TFIIS gene family. Ph.D. Thesis, University of California, Berkeley.

Weber, J., Jelinek, W., and Darnell, J.E., Jr. (1977). The definition of a large viral transcription unit late in Ad2 infection of HeLa cells: mapping of nascent RNA molecules labeled in isolated nuclei. Cell **10**, 611-616.

Weintraub, H., Larsen, A., and Groudine, M (1981). Alpha-Globin-gene switching during the development of chicken embryos: expression and chromosome structure. Cell **24**, 333-344.

Werlen, G., Belin, D., Conne, B., Roche, E., Lew, D.P., and Prentki, M. (1993). Intracellular Ca^{2+} and the regulation of early response gene expression in HL-60 myeloid leukemia cells. J. Biol. Chem. **268**, 16596-16601.

Werner-Washburne, M., Braun, E., Johnston, G.C., and Singer, R.A. (1993). Stationary phase in the yeast *Saccharomyces cerevisiae.* Microbiological Reviews. **57**, 383-401.

West, M.L. and Corden, J.L. (1995). Construction and analysis of yeast RNA polymerase II CTD deletion and substitution mutants. Genetics **140**, 1223-1233.

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

Whitelaw, E. and Proudfoot, N. (1986). Alpha-thalassaemia caused by a poly(A) site mutation reveals that transcriptional termination is linked to 3' end processing in the human alpha 2 globin gene. EMBO J. **5**, 2915-2922.

Wieland, T. and Faulstich, H. (1991). Fifty years of amanitin. Experientia **47**, 1186-1193.

Wiest, D.K. and Hawley, D.K. (1990). *In vitro* analysis of a transcription termination site for RNA polymerase II. Mol. & Cell. Biol. **10**, 5782-5795.

Wiest, D.K., Wang, D., and Hawley, D.K. (1992). Mechanistic studies of transcription arrest at the adenovirus major late attenuation site: Comparison of purified RNA polymerase II and washed elongation complexes. J. Biol. Chem. **267**, 7733-7744.

Williams, L.A. and Kane, C.M. (1995). Isolation and characterization of the *Schizosaccharomyces pombe* gene encoding transcript elongation factor TFIIS. Yeast (in press).

Wolf, D.A., Strobl, L.J., Pullner, A., Eick, D. (1995). Variable pause positions of RNA polymerase II lie proximal to the c-*myc* promoter irrespective of transcriptional activity. Nucleic Acids Res. **23**, 3373-3379.

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

Woychik, N.A., Lane, W.S., and Young, R.A. (1991). Yeast RNA polymerase II subunit RPB9 is essential for growth at temperature extremes. J. Biol. Chem. **266**, 19053-19055.

Woychik, N.A., McKune, K., Lane, W.S., 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. and Young, R. (1989). RNA polymerase II subunit RPB4 is essential for high- and low-temperature yeast cell growth. Mol. & Cell. Biol. **9**, 2854-2959.

Woychik, N.A. and Young, R.A. (1990). RNA polymerase II subunit structure and function. Trends Biochem. Sci. **15**, 347-351.

Woychik, N.A. and Young, R.A. (1994). "Exploring RNA polymerase II structure and function." in Transcription: Mechanisms and Regulation. R.C. Conaway and J.W. Conaway, eds. (New York: Raven Press, Ltd.), pp. 227-242.

Wright, S. and Bishop J.M. (1989). DNA sequences that mediate attenuation of transcription from the mouse protooncogene *myc*. Proc. Natl. Acad. Sci. USA **86**, 505-509.

Wright, S., Mirels, L.F., Calayag, M.C., and Bishop, J.M. (1991). Premature termination of transcription from the P1 promoter of the mouse c-*myc* gene. Proc. Natl. Acad. Sci. USA **88**, 11383-11387.

Wu, J., Ellison, J., Salido, E., Yen, P., Mohandas, T., and Shapiro, L.J. (1994a). Isolation and characterization of XE169, a novel human gene that escapes X-inactivation. Human Molecular Genetics **3**, 153-160.

Wu, J., Salido, E.C., Yen, P.H., Mohandas, T.K., Heng, H.H., Tsui, L.C., Park, J., Chapman, V.M., and Shapiro, L.J. (1994b). The murine Xe169 gene escapes X-inactivation like its human homologue. Nature Genetics **7**, 491-496.

Xu, L., Morgenbesser, S.D., and DePinho, R.A. (1991). Complex transcriptional regulation of *myc* family gene expression in the developing mouse brain and liver. Mol. & Cell. Biol. **11**, 6007-6015.

Xu, M., Barnard, M.B., Rose, S.M., Cockerill, P.N., Huang, S.Y., and Garrard, W.T. (1986). Transcription termination and chromatin structure of the active immunoglobulin kappa gene locus. J. Biol. Chem. **261**, 3838-3845.

Xu, Q., Nakanishi, T., Sekimizu, K., and Natori, S. (1994). Cloning and identification of testis-specific transcription elongation factor S-II. J. Biol. Chem. **269**, 3100-3103.

Yager, T.D. and von Hippel, P.H. (1987). "Transcript elongation and termination in *Escherichia coli*." in *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. F.C. Neidhardt et al., eds. (Washington, D.C.: American Society for Microbiology), pp. 1241-1275.

Yan, D., Weiss, E.A., and Nevins, J.R. (1995). Identification of an activity in B-cell extracts that selectively impairs the formation of an immunoglobulin μ_s poly(A) site processing complex. Mol. & Cell. Biol. **15**, 1901-1906.

Yang, X., Hart, C.M., Grayhack, E.J., and Roberts, J.W. (1987). Transcription antitermination by phage λ gene Q protein requires a DNA segment spanning the RNA start site. Genes & Dev. **1**, 217-226.

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., Yamashita, K., Roy, R., Egly, J., and Bentley, D.L. (1995). The transcriptional elongation inhibitor 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole inhibits transcription factor TFIIH-associated protein kinase. J. Biol. Chem. **270**, 23922-23925.

Yanofsky, C. (1988). Transcription attenuation. J. Biol. Chem. **263**, 609-612.

Yanofsky, C. and Crawford, I.P. (1987). "The tryptophan operon." in *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. F.C. Neidhardt et al., eds. (Washington, D.C.: American Society for Microbiology), pp 1453-1472.

Yarnell, W.S. and Roberts, J.W. (1992). The phage λ gene Q transcription antiterminator binds DNA in the late gene promoter as it modifies RNA polymerase. Cell **69**, 1181-1189.

Ye, Q. and Worman, H.J. (1995). Protein-protein interactions between human nuclear lamins expressed in yeast. Experimental Cell Research **219**, 292-298.

Yoo, O., Yoon, H., Baek, K., Jeon, C., Miyamoto, K., Ueno, A., and Agarwal, K. (1991). Cloning, expression and characterization of the human transcription elongation factor, TFIIS. Nucleic Acids Res. **19**, 1073-1079.

Young, R.A. (1991). RNA polymerase II. Annu. Rev. Biochem. **60**, 689-715.

Yuan, D. and Tucker, P.W. (1984). Transcriptional regulation of the μ-δ heavy chain locus in normal murine B lymphocytes. J. Exp. Med. **160**, 564-583.

Yue, X., Favot, P., Dunn, T.L., Cassady, A.I., and Hume, D.A. (1993). Expression of mRNA encoding the macrophage colony-stimulating factor receptor (c-*fms*) is controlled by a constitutive promoter and tissue-specific transcription elongation. Mol. & Cell. Biol. **13**, 3191-3201.

Zandomeni, R., Zandomeni, M.C., Shugar, D., and Weinmann, R. (1986). Casein kinase type II is involved in the inhibition by 5,6-dichloro-1-β-D-ribofuranosyl-benzimidazole of specific RNA polymerase II transcription. J. Biol. Chem. **261**, 3414-3419.

Zaret, K.S. and Sherman, F. (1982). DNA sequence required for efficient transcription termination in yeast. Cell **28**, 563-573.

Zawel, L., Kumar, P., and Reinberg, D. (1995). Recycling of the general transcription factors during RNA polymerase II transcription. Genes & Dev. **9**, 1479-1490.

Zawel, L. and Reinberg, D. (1995). Common themes in assembly of eukaryotic transcription complexes. Annu. Rev. Biochem. **64**, 533-561.

Zheng, C. and Friedman, D.I. (1994). Reduced Rho-dependent transcription termination permits NusA-independent growth of *Escherichia coli.* Proc. Natl. Acad. Sci. USA **91**, 7543-7547.