RNA Polymerase II Carboxy-Terminal Domain Phosphatase

by

Michael S. Kobor

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

© Copyright by Michael S. Kobor 2001
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

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

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

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

RNA Polymerase II Carboxy-Terminal Domain Phosphatase

© 2001 by Michael S. Kobor. Thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy, Graduate Department of Molecular and Medical Genetics, University of Toronto.

The form of RNA polymerase II (RNAPII) that binds preferentially to promoters is not extensively phosphorylated on the carboxy-terminal heptapeptide repeat domain (CTD) of its largest subunit. The CTD becomes phosphorylated during or shortly after initiation and elongating RNAPII generally has a phosphorylated CTD. Prior to or following transcriptional termination, dephosphorylation of the CTD presumably must occur to regenerate the hypophosphorylated form of RNAPII that is capable of reinitiating transcription. This thesis examines the function of the CTD phosphatase Fcp1p in the yeast Saccharomyces cerevisiae.

In chapter 2, it is shown that Fcp1 is an unusual eukaryotic protein phosphatase that is required for dephosphorylation of the CTD in vivo and for transcription by RNAPII in vivo. These results suggest that Fcp1p is the founding member of a new class of protein phosphatases and acts as a general transcription factor in vivo.

In chapter 3, affinity chromatography is used to study the binding of Fcp1p to TFIIB and the RAP74 subunit of TFIIF. Fcp1p binds in a similar way to both of these factors. RAP74 and TFIIB have a short region of homology and amino acid changes in this region affect the binding to Fcp1p. The genes encoding RAP74 and Fcp1p interact in...
**vivo.** Fcp1p can activate transcription when artificially tethered to a promoter and this effect is largely dependent on binding to RAP74.

In chapter 4, it is shown that yeast strains with mutations in *fcp1* grow much worse when the gene encoding the major CTD kinase Kin28p is also mutated. In contrast, inactivation of another CTD kinase encoded by the *SRB10* gene suppresses the temperature-sensitivity and the sensitivity to certain cell cycle checkpoint inducing drugs of *fcp1* mutant strains. These results therefore suggest that Fcp1p and Srb10p have opposing roles *in vivo.*

In chapter 5, analysis of the phosphorylation state of the CTD reveals that reduced Fcp1p activity results in a increased amount of the largest subunit of RNAPII but this subunit is not incorporated into functional enzyme and is largely degraded at a higher temperature.
Acknowledgments

I would like to express my warmest appreciation for my supervisor Jack Greenblatt. He has been an amazing mentor for whom I have the greatest respect. I learned a tremendous amount about science from Jack and I feel that his guidance prepared my in the best possible way for my own independent career in the academic world.

I also would like to thank the members of my supervisory committee, Paul Sadowski and Jacqueline Segall, for their guidance throughout graduate school. Although not a member of my committee, Jim Ingles always was there to discuss science and life with me and has become a friend. Alan Davidson certainly deserves special recognition and I very much appreciate him for being a great friend who I could discuss science and many other things with. He also introduced me to many aspects of life in North America previously unknown to me.

A large number of people have contributed in multiple ways to the success of my work. Jacques Archambault and Guohua Pan are great scientists and it has been an extremely rewarding experience interacting with them and learning from them. Aside from their scientific mentorship, they also have become great friends. It was a great pleasure for me to teach two summer students, Lisa Simon and Fiona Kouyoumdjian, and to see them effectively completing their tasks. A number other people that are or were associated with the Greenblatt laboratory have been of tremendous help at various points of the work presented in this thesis. These include Nevan Krogan, Bill Lester, James Omichinski, Gouqing Zhong, Paul Jorgensen, and Joyce Li. Similarly, David Jansma, Deming Xu, Chris Koth, Steve Orlicky and Mike Shales were always there when I
needed help or reagents. I have very much enjoyed the company of all of them, the scientific discussions with them and I am grateful for all the assistance and generosity that I received. I also very much appreciate all the friends within the scientific and the rowing community in Toronto that have made me feel very welcome in Canada and made my stay here a so much fun.

Outside of the University of Toronto, I was lucky to have fruitful and successful scientific collaborations with Frank Holstege and Richard Young, Opher Gileadi, Jesper Svjestrup and Steve Buratowski.

My family back home in Germany has always supported and encouraged me in the quest to achieve my goals. To all of them, I am eternally thankful and I am very proud to be related to them. Finally, I would like to thank Manda Ghahremani for her love, her companionship and her support of my work and career. She is all that I could have asked for in a partner.
# Table of Contents

## Chapter 1  
Introduction

1. Basics of transcription by RNA polymerase II (RNAP II)  
2. RNAP II CTD – Importance for enzyme function and transcription  
3. Phosphorylation of the CTD as a central regulatory mechanism  
   3.1 CTD Kinases  
      3.1.1 TFIIH  
      3.1.2 Srb10p/Srb11p cyclin-dependent kinase-cyclin pair  
      3.1.3 Positive Transcription Elongation Factor b (P-TEFb)  
      3.1.4 CTDK-I  
      3.1.5 DNA-dependent protein kinase (DNA-PK)  
      3.1.6 Mitogen-activated protein kinase (MAPK)  
      3.1.7 TIP30/CC3  
      3.1.8 CDC2  
      3.1.9 c-Abl and Abl-related gene product  
3.2 CTD phosphatase  
3.3 Antibodies used to study CTD phosphorylation  
3.4 Complexes involved in initiation of transcription  
   3.4.1 Srb/mediator complex, RNAP II holoenzymes, and the CTD  
   3.4.2 The Srb/mediator complex and transcriptional regulation  
   3.4.3 Other RNAP II complexes  
   3.4.4 Mammalian RNAP II holoenzyme, co-activator complexes and the RNAP II CTD  
   3.4.5 Recruitment of RNAP II as a mechanism of transcriptional activation  
3.5 Transcription Elongation and the CTD  
   3.5.1 TFIIIF  
   3.5.2 CTD phosphatase as an elongation factor  
   3.5.3 DSIF (Spt4p/Spt5p), NELF and P-TEFb
3.5.4 hSPT5, Tat-Sf1, RAP30 and Tat-SF
3.5.5 FACT (Spt16p/Pob3p)
3.5.6 CA150
3.5.7 An elongating form of RNAPII
3.5.8 Linking transcription elongation, DNA repair
    and recombination
3.6 Connecting CTD phosphorylation to mRNA maturation
    3.6.1 Capping
    3.6.2 Splicing and transcription
    3.6.3 Polyadenylation and transcription termination
    3.6.4 Compartmentalization of transcription complexes
3.7 Conformational Changes in the CTD – connecting things
3.8 Physiological changes in RNAPII CTD phosphorylation
    3.8.1 Diauxic shift
    3.8.2 Heat shock
    3.8.3 Cell cycle regulation of transcription and the RNAPII CTD
    3.8.4 The germ-lineage and oocyte fertilization
    3.8.5 Genotoxic agents
4. Ubiquitination of RNAPII
    4.1. Rsp5p ubiquitin ligase and the CTD
    4.2. Ubiquitination and transcription
    4.3. UV radiation and DNA repair
5. Thesis Outline

Chapter 2
An unusual Eukaryotic Protein Phosphatase Required
for Transcription by RNA Polymerase II and CTD
Dephosphorylation in Saccharomyces cerevisiae

Abstract
Introduction
Results
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>A Motif Shared by TFIIF and TFIIB Mediates their Interaction with the RNA Polymerase II CTD Phosphatase Fcp1p in <em>Saccharomyces cerevisiae</em></td>
<td>145</td>
</tr>
<tr>
<td>4</td>
<td>Opposing Effects of the RNA Polymerase II CTD Phosphatase Fcp1p and the CTD Kinase Srb10p in <em>Saccharomyces cerevisiae</em></td>
<td>198</td>
</tr>
<tr>
<td>5</td>
<td>Reduced Activity of the RNA Polymerase II CTD phosphatase Fcp1p Leads to Increased Synthesis of the Largest Subunit of RNA polymerase II in <em>Saccharomyces cerevisiae</em></td>
<td>229</td>
</tr>
</tbody>
</table>
List of Tables

Chapter 1

Table 1. Protein kinases that phosphorylate the CTD 9
Table 2. Subunit composition of human and yeast TFIIH 12
Table 3. Genes encoding subunits of yeast RNAPII associated complexes 33
Table 4. Human co-activator complexes and their relationship to yeast Srb/mediator 52
Table 5. Transcription elongation factors 59
Table 6. CTD-Interacting Proteins 81

Chapter 2

Table 1. Summary of genome wide-expression data 131
List of Figures

Chapter 1

Figure 1. Protein complexes involved in transcription initiation in
Saccharomyces cerevisiae 41
Figure 2. Functional domains of human TFIIF (RAP30/74) 66

Chapter 2

Figure 1. Fcp1p contains a sequence motif essential for it to function
in vivo that is found in a family of small molecule
phosphotransferases and phosphohydrolases 114
Figure 2. Recombinant Fcp1p can hydrolyze p-nitrophenylphosphate
and dephosphorylate the CTD 118
Figure 3. Temperature-sensitive mutations in FCP1 and genetic
interaction with RNAPII 121
Figure 4. Fcp1p is required in vivo for dephosphorylation
of the RNAPII CTD 125
Figure 5. Fcp1p is generally required for transcription by RNAPII in vivo 128

Chapter 3

Figure 1. Domains of Fcp1p required for function in vivo 153
Figure 2. The BRCT domain is essential for mRNA synthesis and CTD
dephosphorylation 156
Figure 3. Fcp1p interacts with the first cyclin-like repeat in TFIIB 160
Figure 4. TFIIB and RAP74 bind in a similar way to Fcp1p 163
Figure 5. A conserved amino acid sequence motif is involved in the
binding of TFIIB and RAP74 to Fcp1p 166
Figure 6. RAP74 (649-735) and RAP74 (649-735) K695E/K699E
are properly folded protein domains

Figure 7. Effects in vivo of mutations in the Fcp1p binding domain of RAP74 169

Figure 8. Transcriptional activation by LexA-Fcp1 fusion proteins 173

Chapter 4

Figure 1. Mutations in the CTD kinases Kin28p and Srb10p have opposite effects on cell growth in yeast strains with mutations in the CTD phosphatase Fcp1p 206

Figure 2. The increased amount of Rpb1p and abnormal CTD phosphorylation pattern present in fcp1 mutant strains is partly reversed by inactivation of Srb10p 210

Figure 3. Sensitivity of fcp1 mutant strains to cell cycle checkpoint-inducing drugs and effects on RNR3 gene expression 213

Figure 4. Reduced Fcp1p activity can partly suppress the effects of the srb10D290A mutation on invasive growth 217

Figure 5. Models for the role of Fcp1p, Srb10p and another CTD kinase in regulating CTD phosphorylation 224

Chapter 5

Figure 1. Depletion of Fcp1p leads to an increased level of Rpb1p 236

Figure 2. Excess Rpb1p subunit is not assembled into RNAPII 240

Figure 3. Free Rpb1p is degraded at 37°C 243

Figure 4. Regions upstream of the translation start codon of RPBI are involved in the increased synthesis of Rpb1p in response to mutations in FCP1 247

Figure 5. Model for the regulation of the translation of the RPB1 mRNA in response to reduced Fcp1p activity 253
Figure 1. Summary of the genetic interactions, physical interactions and catalytic activity involving $FCPI$ described in this thesis
List of Appendices

Plasmids used in this study 351
Yeast Strains used in this study 354
Chapter 1

The Role of the RNA Polymerase II Carboxy-Terminal Domain in mRNA Synthesis
1. Basics of transcription by RNA polymerase II (RNAPII)

In eukaryotic cells, three nuclear RNA polymerases transcribe DNA into RNA. RNAPI is responsible for transcription of ribosomal DNA. RNAPII transcribes protein-coding genes and synthesizes some small nucleolar RNAs, and RNAPIII synthesizes 5S RNA and tRNAs. The three enzymes have multi-subunit structures that are partially conserved among all cellular RNA polymerases. The two largest subunits are homologous the β and β’ subunits of *Escherichia coli* RNAP and are associated with a number of smaller subunits. Some of the smaller subunits are common to all three eukaryotic RNAPs, whereas others are shared only by RNAPI and RNAPIII (Sawadogo and Sentenac 1990; Young 1991).

The purified RNAPs are unable to initiate transcription specifically at promoters. In the case of RNAPII, specific transcription initiation requires the assembly at the core promoter of a preinitiation complex (PIC), consisting of RNAPII and the general transcription factors (GTFs). Many protein-DNA and protein-protein interactions occur within the transcription initiation complex (Orphanides et al. 1996). First, the transcription factor TFIID binds to promoter DNA through a direct interaction of its TBP (TATA-binding protein) subunit with the TATA element, located approximately 30 nucleotides upstream of the transcription start sites in the majority of protein-coding genes. TBP-associated factors (TAFs) and other promoter elements, the initiator and downstream promoter elements, participate in this interaction as well. TFIIA is able to enhance the interaction of TFIID with promoter. TFIIB interacts with TBP and the
promoter DNA and helps to recruit RNAPII along with TFIIF (RAP30/RAP74). This provides the scaffold for the addition of TFIIE and TFIIH, which complete the preinitiation complex (Buratowski 2000). Localized melting of the duplex promoter DNA by an ATP-dependent helicase within TFIIH leads to the formation of an open complex. This is then followed by formation of the first phosphodiester bond and promoter clearance by RNAPII. However, in many cases, abortive initiation precedes the escape of RNAPII into processive elongation (Conaway et al. 2000). Transcription elongation then continues until termination signals are reached.

The largest subunit of RNAPII has a unique carboxy-terminal domain (CTD) comprised of tandem repeats of the consensus sequence Tyr-Scr-Pro-Thr-Ser-Pro-Ser (YSPTSPS). In this introduction, I will focus on the role of the RNAPII CTD in the production of mRNAs. I will first describe the basic properties of the CTD and the proteins involved in phosphorylation of the CTD, its major reversible modification. A section about RNAPII holoenzyme complexes and co-activator complexes that participate in gene activation and their connection to the CTD then follows. Next, a section on transcription elongation provides insight into complex regulatory mechanisms involving the CTD. I will then outline recent studies that suggest that the CTD functions at the interface of transcription and mRNA processing. It is possible that conformational changes of the CTD might connect transcription elongation and termination and this is described in the next section. The focus on CTD phosphorylation then concludes with a description of physiological changes that occur to the CTD in response to intracellular or extracellular stimuli. This will lead to a section about another CTD modification, namely
ubiquitination. Finally, I will relate the experimental work presented in this thesis to the issues discussed in this introduction.

2. RNAPII CTD – Importance for enzyme function and transcription

The YSPTSPS heptapeptide repeat consensus sequence has been conserved through evolution, although the number of the repeats varies among species. RNAPII in the yeast *Saccharomyces cerevisiae* has 25/26 repeats and RNAPII in *Drosophila melanogaster* has 42 repeats, whereas the human and mouse RNAPIIs have 52 repeats (Allison et al. 1985; Corden et al. 1985). In mammals, the CTD consists of mostly perfect consensus repeats in its amino-terminal half but has many non-consensus heptapeptide repeats in its carboxy-terminal half, mostly involving amino acid substitutions at position seven (Dahmus 1996).

The importance of the CTD for enzyme function *in vivo* was established shortly after its discovery. Shortening of the CTD from 26 to 8 repeats in yeast results in a lethal phenotype. Yeast strains in which the CTD is shortened to 11-19 repeats are cold-sensitive, indicative of a defect in enzyme function (Nonet et al. 1987; Allison et al. 1988). Similarly, deletion analysis in *Drosophila* and human tissue culture cells revealed that the CTD could not be completely deleted without severely affecting the viability of the cells. In human cells, deletion of up to 28 repeats does not affect viability, whereas RNAPII molecules with 5 CTD repeats do not support long-term growth (Bartolomei et al. 1988). Recently, a mouse knockout study has confirmed that a certain length of the CTD is necessary for development. Mice with shortened CTDs are smaller than their
wild type littermates and many of them die \textit{in utero} (Litingtung et al. 1999).

Activated transcription driven by the upstream activating sequence (UAS) elements of the \textit{INO1} and \textit{GAL10} promoters is defective in yeast strains carrying RNAPII with a shortened CTD but basal transcription from these promoters is not affected (Scafe et al. 1990). In contrast, the \textit{HIS4} gene is activated to wild-type levels in a CTD truncation strain (Scafe et al. 1990). Lengthening the CTD can suppress mutations that weaken activation by the yeast activator protein Gal4p \textit{in vivo} (Allison and Ingles 1989). In rich medium, truncation of the CTD preferentially reduces the expression of genes encoding glycolytic enzymes and this is dependent on the promoters of these genes (Meisels et al. 1995). The integrity of the CTD also is required for enhancer-driven transcription in tissue culture cells but not for transcription regulated by Sp1, a factor that activates from promoter-proximal positions (Gerber et al. 1995). However, recent experiments in yeast indicate that a small number of genes, including \textit{CUP1}, \textit{HSP82} and \textit{SSA4}, can be transcribed \textit{in vivo} by RNAPII molecules that lack the CTD (McNeil et al. 1998).

In nuclear extract transcription systems derived from yeast, progressive truncation of the CTD causes progressive loss of transcription in response to the activator Gal4-Vp16p as well as defects in factor-independent transcription, but has little effect on transcription elongation or termination (Liao et al. 1991). Similarly, CTD-specific monoclonal antibodies can inhibit transcription initiation but not elongation in mammalian \textit{in vitro} transcription systems (Laybourn and Dahmus 1989; Moyle et al. 1989; Thompson et al. 1989)
At some promoters such as the dehydrofolate reductase promoter, a promoter that lacks a TATA box, RNAPIIB, a form of the enzyme that lacks the CTD, does not assemble into the preinitiation complex (Kang and Dahmus 1993). Consistent with this, \textit{in vitro} transcription from this promoter is dependent on the CTD (Kang and Dahmus 1993). In contrast, transcription from the TATA box-containing adenovirus-2 major late promoter is independent of the CTD \textit{in vitro} (Kang and Dahmus 1993). Certain core promoter elements including the TATA box play a role in determining the CTD dependence for transcription of a gene \textit{in vitro} (Buermeyer et al. 1995). In certain highly purified systems, the CTD is also dispensable for activated transcription from a TATA box-containing promoter (Sun et al. 1998).

Thus, these studies show that CTD truncation can cause specific defects in uninduced as well as induced transcription \textit{in vivo} and \textit{in vitro} but the molecular understanding of CTD function relied on the discovery of the SRB/mediator and the observation that the CTD is modified \textit{in vivo}. Phosphorylation is the most prominent post-transcriptional modification of the CTD and I will discuss this in some detail, but glycosylation (Kelly et al. 1993) and ubiquitination (Bregman et al. 2000) also occur.

3. Phosphorylation of the CTD as a central regulatory mechanism

Two major forms of RNAPII that are distinguished by their degree of CTD phosphorylation exist \textit{in vivo} (Cadena and Dahmus 1987). The hyperphosphorylated form of the enzyme has been designated RNAPIIO, whereas the hypophosphorylated form is called RNAPIIA. RNAPIIo and RNAPIIa refer to the hypophosphorylated and
hyperphosphorylated form of the largest subunit of RNAPII (Rpb1p). I will use this nomenclature throughout this thesis. A number of intermediately phosphorylated forms have also been reported. The CTD is phosphorylated in vivo predominantly on serine residues but a low level of phosphorylation also occurs on threonine and tyrosine (Cadena and Dahmus 1987; Zhang and Corden 1991; Baskaran et al. 1993). Although the exact number of sites phosphorylated in vivo has not been quantitated, it appears that human RNAPII carries more than 50 phosphates (Dahmus 1994). Ser2 or Ser5 of the CTD heptapeptide repeat YSPTSPS cannot be substituted by Ala or Glu in yeast, suggesting an essential function for CTD phosphorylation in vivo (West and Corden 1995). Changing Tyr1 to Phe also results in lethality (West and Corden 1995).

Phosphorylation of the CTD accompanies mRNA synthesis. The hypophosphorylated form of RNAPII is preferentially recruited to the transcription initiation complex (Lu et al. 1991; Chesnut et al. 1992). It is also possible that protein-protein interactions mediated by the unphosphorylated CTD play a role in the positioning of RNAPII at the core promoter. The unphosphorylated CTD can interact directly with TBP (Usheva et al. 1992) and also might interact with the RAP74 subunit of TFIIF and the small subunit of TFIIE (Kang and Dahmus 1995). Transcription from at least some promoters in highly defined in vitro systems is not dependent on CTD phosphorylation, whereas transcription in less defined systems requires CTD phosphorylation (Li and Kornberg 1994).

In general, transition of the transcription complex from initiation to stable elongation is accompanied by hyperphosphorylation of the CTD (Dahmus 1996).
Transcription complexes paused near the transcriptional start site on a number of
*Drosophila* genes contain RNAPIIA (O'Brien et al. 1994). Induction of transcription and
the release of RNAPII from the paused complex correlates with phosphorylation of the
CTD (O'Brien et al. 1994). The hyperphosphorylated RNAPIIO is the prominent form
of the polymerase at sites of highly productive elongation such as large developmental
puffs (Weeks et al. 1993). However, hyperphosphorylated and hypophosphorylated
RNAPII work equally well in non-specific elongation assays *in vitro* (Yamaguchi et al.
1999a).

It is emerging that phosphorylation of the CTD at multiple sites might serve to
disrupt interactions between the unmodified CTD and proteins necessary for the
formation of a stable preinitiation complex. As well, the phosphorylated CTD provides
a binding platform for factors involved in later steps of mRNA synthesis (Bentley 1999;
Hirose and Manley 2000; Proudfoot 2000). A simple model for the periodic cycling of
the phosphorylation state of RNAPII during transcription has been suggested that
involves CTD kinases that act during transcription initiation and elongation and a CTD
phosphatase that regenerates unphosphorylated RNAPIIA during or after transcription
termination (Dahmus 1996).

### 3.1 CTD Kinases

Various protein kinases can phosphorylate the CTD of RNAPII or synthetic
peptides containing the consensus CTD repeat sequence *in vitro* and the known CTD
kinases are listed in Table 1.
Table 1. Kinases that phosphorylate the RNAPII CTD

<table>
<thead>
<tr>
<th>Kinase</th>
<th>Interacting proteins</th>
<th>Phosphorylated CTD residue</th>
<th>Function in transcription</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDK7</td>
<td>CyclinH</td>
<td>Ser5 \textit{in vitro} and \textit{in vivo}; Some Ser2 reported</td>
<td>Positive factor, generally required for transcription \textit{in vivo}; CTD kinase activity dispensable for \textit{in vitro} transcription in purified systems.</td>
</tr>
<tr>
<td>Kin28p</td>
<td>Ccl1p (subunits of TFIIH, see Table 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDK8</td>
<td>CyclinC</td>
<td>Ser5 \textit{in vitro} and Ser2 \textit{in vitro}</td>
<td>Negative regulator of transcription for 173 genes in yeast; CTD phosphorylation by Srb10p prevents formation of the PIC; positive role in activation by Gal4p.</td>
</tr>
<tr>
<td>Kin28p</td>
<td>Srb11p (found associated with yeast RNAPII holoenzyme and some mammalian mediator complexes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDK9</td>
<td>CyclinT1 CyclinT2a CyclinT2b CyclinK</td>
<td>Ser2 \textit{in vitro}; Ser2 and Ser5 \textit{in vitro} in the presence of HIV Tat</td>
<td>Positive transcription elongation factor, required for activation by HIV Tat.</td>
</tr>
<tr>
<td>Ctklp</td>
<td>Ctk2p Ctk3p</td>
<td>Ser2 transiently during diauxic shift \textit{in vivo}; possible negative regulation of Ser5 during</td>
<td>Ctklp and Ctk2p have homology to CDK9 and cyclinT; might have a role in activation and</td>
</tr>
<tr>
<td>Ctk2p</td>
<td>Ctk3p</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ctk3p</td>
<td>Ctk2p</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>Logarithmic Growth</td>
<td>Repression of Transcription</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-------------------</td>
<td>-----------------------------</td>
<td></td>
</tr>
<tr>
<td>DNA-PK</td>
<td>Ku (DNA binding subunit)</td>
<td>Ser2 and Ser7 \textit{in vitro}</td>
<td>Might have a role in multiple round transcription.</td>
</tr>
<tr>
<td>MAP kinases</td>
<td>Ser5 \textit{in vitro}</td>
<td>Phosphorylate the CTD during heat shock and upon fertilization.</td>
<td></td>
</tr>
<tr>
<td>TIP30/CC3</td>
<td>N.D.</td>
<td>Enhances HIV Tat-dependent transcription; upregulation of some apoptotic genes.</td>
<td></td>
</tr>
<tr>
<td>CDC2</td>
<td>CyclinB</td>
<td>Ser2 and Ser5 \textit{in vitro}</td>
<td>Phosphorylation of RNAPII by Cdc2p \textit{in vitro} can lead to inhibition of transcription</td>
</tr>
<tr>
<td>c-Abl Arg</td>
<td>Tyr1 \textit{in vitro and in vivo}</td>
<td>Possible effect on transcription initiation and elongation.</td>
<td></td>
</tr>
</tbody>
</table>
3.1.1 TFIID

TFIID is a multisubunit general transcription factor with associated CTD kinase and DNA helicase activities (Coin and Egly 1998). It can be resolved into a core complex, containing 5 or 6 subunits, and a 3-subunit complex containing the kinase activity. The genes for the human and yeast subunits are homologous to a large extent but slight differences exist in the composition of the core TFIID. The genes encoding the TFIID subunits are essential for viability in yeast. Table 2 lists the yeast and human TFIID subunits and summarizes their properties.
Table 2. Subunit composition of yeast and human TFIIH

<table>
<thead>
<tr>
<th>Yeast subunit</th>
<th>Size (kDa)</th>
<th>Human homolog</th>
<th>Core subunits yeast/human</th>
<th>Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ssl2p</td>
<td>95</td>
<td>XPB</td>
<td>- / +</td>
<td>3’-5’ DNA helicase, ATPase; essential for promoter opening and promoter escape.</td>
</tr>
<tr>
<td>Rad3p</td>
<td>89</td>
<td>XPD</td>
<td>+/-sometimes</td>
<td>5’-3’ DNA helicase, ATPase, required for DNA repair.</td>
</tr>
<tr>
<td>Tfb1p</td>
<td>73</td>
<td>p62</td>
<td>+ / +</td>
<td>Required for nucleotide excision repair, target for activators.</td>
</tr>
<tr>
<td>Tfb2p</td>
<td>59</td>
<td>p52</td>
<td>+ / +</td>
<td>Required for nucleotide excision repair.</td>
</tr>
<tr>
<td>Ssl1p</td>
<td>52</td>
<td>p44</td>
<td>+ / +</td>
<td>Required for nucleotide excision repair; involved in DNA binding and stimulation of XPD activity.</td>
</tr>
<tr>
<td>Ccl1p</td>
<td>45</td>
<td>CyclinH</td>
<td>- / -</td>
<td>Cyclin subunit of cyclin-dependent CTD kinase.</td>
</tr>
<tr>
<td>Tfb3p</td>
<td>38</td>
<td>MAT1</td>
<td>+ / -</td>
<td>Required for nucleotide excision repair; MAT1/CDK7/CyclinH form the CAK subcomplex (human).</td>
</tr>
<tr>
<td>Tfb4p</td>
<td>37</td>
<td>p34</td>
<td>- / +</td>
<td>Required for nucleotide excision repair.</td>
</tr>
<tr>
<td>Kin28p</td>
<td>35</td>
<td>CDK7</td>
<td>- / -</td>
<td>Kinase subunit of cyclin-dependent CTD kinase; Kin28p and Ccl1p form the TFIIK subcomplex (yeast).</td>
</tr>
</tbody>
</table>
The mammalian Cdk-activating kinase (CAK) complex consists of CDK7, cyclinH and MAT1, all of which are subunits of TFIIH. This three-subunit CAK complex phosphorylates various cyclin-dependent kinases but the RNAPII CTD is the predominant substrate when CAK is associated with core TFIIH (Rossignol et al. 1997; Yankulov and Bentley 1997). Although TFIIH can specifically phosphorylate the CTD of RNAPII in the absence of any other factors, TFIIH kinase activity is greatly stimulated by factors promoting the assembly of RNAPII into the preinitiation complex and by DNA elements capable of directing the formation of a transcription competent complex (Lu et al. 1992; Hengartner et al. 1998).

TFIIH phosphorylates the CTD \textit{in vitro} predominantly on Ser5 (Roy et al. 1994; Trigon et al. 1998) but might also phosphorylate Ser2 (Dubois et al. 1997). The kinase activity is proline-directed because the replacement of either Pro by Ala within a heptapeptide repeat leads to abolition of peptide phosphorylation (Trigon et al. 1998). TFIIH has a 13-fold better kinase activity on a CTD-2 peptide compared to a CTD-1 peptide (Trigon et al. 1998)(number indicates how many times the heptapeptide is repeated) and generally prefers longer CTD-like substrates (Rickert et al. 1999). Also, kinase activity is strongly enhanced by replacing Thr4 with Leu and Ser7 with Lys in a CTD-4 peptide (Rickert et al. 1999). CDK7/cyclinH/MAT1 phosphorylates the amino-terminal, more conserved half of the CTD to the same extent as it phosphorylates the full-length mouse CTD, but displays a fourfold preference for the carboxy-terminal half of the CTD that contains a high percentage of non-consensus repeats (Rickert et al. 1999). This suggests that TFIIH can differentially phosphorylate different portions of
the CTD. In addition, TFIH can phosphorylate other components of the general transcription machinery besides the CTD of RNAPII such as TFIIE, TFIIF and TBP (Ohkuma and Roeder 1994) as well as the coactivator PC4 (Malik et al. 1998).

A number of factors can regulate the CTD kinase activity of TFIH. TFIIE enters the preinitiation complex together with TFIH and can stimulate the TFIH DNA helicase and CTD kinase activities (Ohkuma and Roeder 1994). The mediator complex (described below) involved in transcriptional activation can also stimulate the CTD kinase activity of yeast TFIH (Kim et al. 1994). Many transcription activator proteins such as HSV VP16, human p53, and HIV Tat can bind TFIH directly (Xiao et al. 1994; Blau et al. 1996; Parada and Roeder 1996) and also might stimulate TFIH kinase activity (Parada and Roeder 1996). Conversely, the transcriptional coactivator PC4 (Schang et al. 2000) and the tumor suppressor p16INK4A (Serizawa 1998) can act as inhibitors of CDK7-mediated CTD phosphorylation. It has also been suggested that hnRNP U can inhibit TFIH CTD kinase activity within the RNAPII holoenzyme, and this results in a defect in transcriptional elongation (Kim and Nikodem 1999). The TFIH CTD kinase and transcription activities might also be negatively regulated by CDK8/cyclinC (described below) through phosphorylation of cyclinH (Akoulitchev et al. 2000).

Compelling evidence for a physiological role of the TFIH-associated CTD kinase activity was obtained as a result of using a conditional temperature-sensitive allele of the KIN28 gene in S. cerevisiae. Upon shift to the non-permissive temperature, RNAPII molecules with a hypophosphorylated CTD accumulate and those with a hyperphosphorylated CTD disappear (Valay et al. 1995). These results suggest that
Kin28p is an important CTD kinase *in vivo*. Moreover, the synthesis of poly(A)^+ mRNA is significantly reduced in this mutant strain at the non-permissive temperature (Valay et al. 1995). These results were confirmed by genome-wide expression analysis with DNA microarrays. The mRNA levels for almost all genes were reduced by more than 2-fold in a *kin28* strain after the cells were kept for 45 min at the non-permissive temperature (Holstege et al. 1998). Importantly, the mRNA levels for 94% of all yeast genes are as dependent on the TFIIIH kinase as they are on RNAPII itself (Holstege et al. 1998). However, the transcription of a small number of genes, including *CUP1*, *SSA4* and *HSP82*, can be highly induced *in vivo* in the absence of Kin28p activity (Lee and Lis 1998; McNeil et al. 1998). Consistent with the *in vitro* phosphorylation data, TFIIIH seems to phosphorylate Ser5 of the YSPTSPS heptapeptide repeat *in vivo* and, interestingly, Ser5 phosphorylation by TFIIIH appears confined to promoter regions (Komarnitsky et al. 2000).

However, it is unclear whether TFIIIH CTD kinase activity is needed for the actual transcription process. The requirement for Kin28p in yeast may be for the Kin28 protein rather than its kinase activity. Moreover, transcription in purified systems can occur in the absence of CTD phosphorylation or the CDK7 kinase activity, although CDK7 seems to stimulate the production of long run-off transcripts (Akoulitchev et al. 1995; Makela et al. 1995). Most convincingly, transcription initiation *in vitro* using recombinant TFIIIH subcomplexes can occur in the absence of CAK (Tirode et al. 1999). Although CAK can stimulate *in vitro* transcription in conjunction with XPD, this is independent of the CTD kinase activity of CDK7 (Tirode et al. 1999).
A number of studies indicate that promoter opening is the essential function of TFIIH in transcription initiation. The ATP-dependent DNA helicase activity of the XPB/Ssl2p subunit of TFIIH is required for promoter melting in vivo (Guzman and Lis 1999) and in vitro (Tirode et al. 1999), likely reflecting the ATP-cofactor requirement for transcription initiation. Transcription initiation with recombinant TFIIH absolutely depends on XPB DNA helicase activity (Tirode et al. 1999), and a yeast strain with a conditional lethal mutation in the ATP-binding site of Ssl2p, the homolog of XPB, is defective in RNAPII transcription (Qiu et al. 1993; Guzder et al. 1994). The XPB DNA helicase of TFIIH also might be required for promoter escape and early elongation in a TFIIE-dependent manner (Dvir et al. 1997; Kugel and Goodrich 1998; Kumar et al. 1998; Moreland et al. 1999; Bradsher et al. 2000). The XPD helicase is dispensable but stimulates transcription and allows the CAK complex to be anchored to TFIIH (Tirode et al. 1999). Most importantly, CDK7-mediated phosphorylation of the CTD can occur in the absence of promoter opening, independently of the essential TFIIH DNA helicase (Tirode et al. 1999). In any case, although Kin28p activity is clearly required for mRNA synthesis in vivo, as most clearly demonstrated in the yeast system, this might be mainly due to an effect on mRNA processing events that are connected to transcription through the RNAPII CTD (Hirose and Manley 2000). I will discuss the evidence for this below.

In addition to transcription, TFIIH also has essential functions in nucleotide excision repair (NER) and cell cycle regulation (Frit et al. 1999). Only the core subunits are necessary for TFIIH to function in NER, whereas the kinase subcomplex is not required (Svejstrup et al. 1995). The DNA helicase activity of Rad3p/XPB, which is not
required for transcription, seems to play a crucial role in NER (Sung et al. 1988; Feaver et al. 1993).

Mammalian CDK-activating kinase (CAK) is essential to phosphorylate and thereby activate CDKs involved in cell-cycle regulation. This function of mammalian CAK is independent of its role in TFIIH (Rossignol et al. 1997; Yankulov and Bentley 1997). In yeast, these functions are split: Kin28p functions only as the kinase subunit of TFIIH (Cismowski et al. 1995), whereas Cak1p functions as a CAK (Kaldis et al. 1996). However, Cak1p can phosphorylate Kin28p, which is itself a CDK, on Thr 162, and this stimulates the CTD kinase activity of Kin28p (Kimmelman et al. 1999). Therefore, although mammalian CDK7 and yeast Cak1p are different forms of CAK, both control the cell cycle and the phosphorylation of the RNAPII CTD by TFIIH.

Understanding of TFIIH will be greatly aided by recent structural studies. Human TFIIH forms a ring-like structure from which protrude a large protein domain containing CDK7, CyclinH and MAT1 (Schultz et al. 2000). P44 is at the base of the protruding domain and is flanked by the XPB and XPD helicases (Schultz et al. 2000). The core 5-subunit form of yeast TFIIH also forms a ring structure consisting of the Tfb1p, Tfb2p and Tfb3p subunits to which the Ssl2p and Rad3p polypeptides are appended (Chang and Kornberg 2000).

3.1.2 Srb10p/Srb11p cyclin-dependent kinase-cyclin pair

Yeast strains carrying certain CTD deletion alleles have cold-sensitive phenotypes. A genetic screen for suppressors of this phenotype led to the identification
of the *SRB* genes (Suppressor of RNAP B) (Nonet and Young 1989). *SRB10* and *SRB11* are non-essential genes encoding another cyclin-dependent kinase/cyclin pair within the transcription apparatus (Liao et al. 1995). Srb10p is related to the CDC2 subfamily of serine/threonine protein kinases, and the recombinant Srb10p/Srb11p complex phosphorylates the CTD *in vitro* on Ser5 (Hengartner et al. 1998). The homologous human CDK8/cyclinC complex can also phosphorylate the CTD on Ser5, although phosphorylation at Ser2 also has been reported (Sun et al. 1998; Rickert et al. 1999). The CDK8/cyclinC complex phosphorylates recombinant mouse GST-CTD equally well on the amino-terminal, highly conserved half and the carboxy-terminal half, which has a larger number of non-consensus repeats (Rickert et al. 1999). Amino acid substitutions of the last residue of the consensus heptapeptide repeat do not affect phosphorylation of Ser5 by CDK8/cyclinC (Rickert et al. 1999). Furthermore, unlike CDK7/cyclinH, CDK8/cyclinC preferentially phosphorylates shorter CTD-like substrates (Rickert et al. 1999). CDK8/cyclinC and CDK7/cyclinH/MAT1 may also differentially phosphorylate the CTD *in vivo* (Rickert et al. 1999). The CTD kinase activity of these two complexes is differentially inhibited *in vitro* by the ATP analogs H7 (1-(5-isoquinolinoylsulfonyl)-3-methylpiperazine) and H8 (N-[2-(methylamino)-ethyl]-5-isoquinoline sulfonamide), whereas DRB (5,6-dichloro-1-b-M-ribofuranosylbenzimidazole) inhibits the two CTD kinases to the same extent (Rickert et al. 1999).

Srb10p and Srb11p are components of some purified yeast RNAPII holoenzyme preparations (see below) (Liao et al. 1995). Supporting a role for Srb10p in CTD phosphorylation, RNAPII holoenzymes purified from yeast strain with a mutation in the
ATP-binding site of Srb10p have a 5-10 fold lower level of CTD phosphorylation after incubation with ATP in comparison with the same enzyme purified from wild-type cells after incubation with ATP (Liao et al. 1995). However, RNAPII holoenzyme lacking Srb10p does not have a defect in basal or activated transcription (Liao et al. 1995).

Genome-wide expression studies revealed that Srb10p is a negative regulator of transcription in vivo, since 173 genes are constitutively de-repressed in a srb10 strain (Holstege et al. 1998). This is supported by the finding that the SRB10 and SRB11 genes, also known as SSN3 and SSN8, are essential for complete repression of the GAL genes involved in galactose metabolism (Carlson 1997). Additionally, Srb10p and Srb11p are required for transcriptional repression by Tup1p/Ssn6p, a common corepressor in yeast (Carlson 1997). I will discuss the likely reason for the negative role of Srb10p below. However, Srb10p also has a positive role in GAL gene activation, probably mediated by phosphorylation of the transcriptional activator Gal4p within the context of the RNAPII holoenzyme (Hirst et al. 1999).

Lastly, human CDK8/cyclinC also has a negative role in transcription. CDK8 and cyclinC are components of several coactivator complexes involved in transcriptional regulation (Malik and Roeder 2000) (see below). They may phosphorylate TFIIH, leading to inhibition of the TFIIH CTD kinase activity and reduced transcriptional activation (Akoulitchev et al. 2000).

3.1.3 Positive Transcription Elongation Factor b (P-TEFb)

DRB (5,6-dichloro-1-β-M-ribofuranosylbenzimidazole) is a protein kinase
inhibitor that blocks CTD phosphorylation and induces arrest of elongating RNAPII in vivo (Sehgal et al. 1976; Tamm et al. 1980) and in crude transcription systems in vitro (Chodosh et al. 1989; Zandomeni 1989; Marshall and Price 1992). P-TEFb was purified as a positively acting, DRB-sensitive factor that can overcome transcriptional arrest by RNAPII complexes during early elongation (Marshall and Price 1995). P-TEFb can phosphorylate RNAPII in these complexes but also efficiently phosphorylates free RNAPII (Marshall et al. 1996). However, extensive phosphorylation of the CTD by P-TEFb in a processive manner occurs only when the CTD is already partly phosphorylated (Marshall et al. 1996). DRB-sensitive transcripts only arise when transcription reactions are performed with RNAPIIA containing the CTD and not with RNAPIIIB that lacks the CTD (Marshall et al. 1996). This suggests that the CTD is an important target for P-TEFb.

Native human P-TEFb consists of the cyclin-dependent kinase CDK9, a regulatory cyclin subunit and potentially several other polypeptides (Zhu et al. 1997; Peng et al. 1998a; Peng et al. 1998b). The cyclin partner is required for CTD kinase activity and stimulation of transcription (Peng et al. 1998a). Three different cyclins, cyclinT1, cyclinT2a and cyclinT2b, can regulate CDK9 (Peng et al. 1998b) and recent findings suggest that CDK9 can also associate with cyclinK in vivo (Edwards et al. 1998). CyclinK can interact with RNAPII, and immunoprecipitated cyclinK containing complexes have CTD kinase and CDK activity (Edwards et al. 1998). A CDK9-cyclin K complex can phosphorylate the CTD in vitro and can functionally substitute for P-TEFb in transcription reactions in vitro (Fu et al. 1999).
P-TEFb activity has also been shown to be important for HIV Tat mediated activation *in vivo* and *in vitro* (Mancebo et al. 1997; Yang et al. 1997; Zhu et al. 1997). P-TEFb is apparently recruited to HIV transcription complexes through interactions of Tat with both cyclinT1 and HIV RNA (Bieniasz et al. 1998; Wimmer et al. 1999; Taube et al. 2000). The kinase activity of P-TEFb is required for its positive effect on HIV Tat-mediated stimulation of transcription elongation and, importantly, there is a close correlation between the inhibition of P-TEFb CTD kinase activity and transcriptional activation by Tat for a large panel of drugs (Mancebo et al. 1997). It has been suggested that CDK9 phosphorylates Ser2 of the CTD heptapeptide repeat, but the substrate specificity is altered in the presence of HIV Tat such that CDK9 phosphorylates both Ser2 and Ser5 (Zhou et al. 2000a). CIITA is a transcriptional activator for class II major histocompatibility (MHC II) genes that can bind P-TEFb *in vitro* (Kanazawa et al. 2000). Interestingly, HIV Tat competes with CIITA for binding to P-TEFb and blocks the expression of MHC class II genes (Kanazawa et al. 2000).

On *Drosophila* polytene chromosomes, P-TEFb is located at >200 distinct sites. After heat shock it is rapidly recruited to heat shock loci in a heat shock factor-dependent manner (Lis et al. 2000). P-TEFb also frequently co-localizes with hypophosphorylated RNAPIIA at promoter pause sites, supporting a model in which P-TEFb acts to stimulate promoter-paused RNAPII to enter into productive elongation (Lis et al. 2000). As discussed below, P-TEFb appears to stimulate early elongation as a result of its ability to overcome the effects of the negative elongation factors DSIF and NELF.
3.1.4. CTDK-I

CTDK-I is another CDK that was purified from yeast and specifically phosphorylates the CTD in vitro (Lee and Greenleaf 1989). CTDK-I consists of three subunits encoded by CTK1, CTK2 and CTK3, all of which are non-essential genes (Lee and Greenleaf 1991; Sterner et al. 1995). Ctk1p is the catalytic subunit and its primary amino acid sequence resembles that of the CDK9 subunit of P-TEFb (Lee and Greenleaf 1991). Ctk2p is closely related to cyclinT, whereas Ctk3p has no homology to known proteins (Sterner et al. 1995). Like P-TEFb, CTDK-I can stimulate elongation by RNAPII in vitro, although this has been shown only in assays that used human transcription systems (Lee and Greenleaf 1997). In contrast to TFIIH and Srb10p/Srb11p, the presence of CTDK-I has not been reported in any of the purified yeast RNAPII holoenzyme complexes (Sterner et al. 1995).

Deletions of ctk1, ctk2 and ctk3 in yeast cause slow growth, cold sensitivity, flocculence and an enlarged cell size phenotype (Lee and Greenleaf 1991; Sterner et al. 1995). Importantly, a ctk1Δ strain has an abnormal CTD phosphorylation pattern with an increase of RNAPIIa and a decrease of RNAPIIo (Lee and Greenleaf 1991). Some phosphorylation remains in this strain, consistent with the existence of other physiological CTD kinases (Lee and Greenleaf 1991). Ctk1p might transiently phosphorylate Ser2 of the CTD during diauxic shift, leading to the induction of genes encoding glycogen synthase (GSY2) and cytosolic catalase (CTTI) (Patturajan et al. 1999). In logarithmically growing cells, Ctk1p might have a negative effect on Ser5...
phosphorylation. Expression of the SSA3 and SSA4 genes is increased in the ctk1Δ strain, suggesting a possible role for CTDK-I in transcriptional repression (Patturajan et al. 1999), although it is not possible in these experiments to distinguish direct and indirect effects of CTDK-I.

3.1.5 DNA-dependent protein kinase (DNA-PK)

DNA-PK is a large, multi-subunit complex that consists of a 350kDa catalytic subunit and a DNA binding component that corresponds to the Ku autoantigen. DNA-PK can phosphorylate the CTD of RNAPII (Dvir et al. 1993) and many other proteins, including TBP, TFIIB (Chibazakura et al. 1997), Hsp90, p53 and c-Jun. DNA-PK-mediated phosphorylation of the CTD is dependent on DNA and on the assembly of transcription initiation complexes containing the general transcription factors TFIID, TFIIB and TFIIF (Dvir et al. 1992). DNA-PK phosphorylates Ser2 and Ser7, but not Ser5, of the YSPTSPS heptapeptide repeat, and the amino acids surrounding the phosphorylated residues are not absolutely necessary for this phosphorylation (Trigon et al. 1998). Thr4 of the YSPTSPS heptapeptide repeat is not a substrate for DNA-PK, but replacing Thr4 with Glu results in a dramatic increase of Ser2 phosphorylation by DNA-PK (Trigon et al. 1998). In contrast, a replacement of Tyr1 with Ala leads to a decrease in CTD peptide phosphorylation, possibly indicating a role for this residue in enzyme-substrate binding (Trigon et al. 1998).

Nuclear extracts from DNA-PK-deficient cells have a two-sevenfold decrease in
the level of *in vitro* transcription from several promoters when compared with extracts from wild-type cells. This may result from a reduced ability of the extract lacking DNA-PK to perform multiple round transcription (Woodard et al. 1999), but no connection to the CTD was shown. Certain transcriptional activator proteins, such as heat shock factor, can increase the CTD kinase activity of DNA-PK (Peterson et al. 1995).

3.1.6 Mitogen-activated protein kinase (MAPK)

The MAP kinases ERK1 and ERK2 phosphorylate Ser5 of the CTD heptapeptide repeat with no apparent preference for consensus versus non-consensus repeats (Trigon et al. 1998). There is a small and progressive increase in the phosphorylation of peptides containing one, two or three CTD motifs (Trigon et al. 1998). Like TFIIH, the MAP kinases are proline-directed CTD kinases whose ability to phosphorylate CTD peptides depends on the presence of Pro residues (Trigon et al. 1998). In contrast, replacing Thr4 by Leu enhances CTD peptide phosphorylation by ERK1/ERK2.

MAP kinases are induced by mitogenic stimuli and heat shock, and this may result in phosphorylation of the CTD by ERK1/2 *in vivo* (see below).

3.1.7 TIP30/CC3

TIP30/CC3 is a novel serine/threonine kinase that can phosphorylate a recombinant GST-CTD fusion protein on both serine and threonine residues (Xiao et al. 2000). CTD phosphorylation is dependent on a functional putative ATP-binding motif in
TIP30 and can be stimulated by the HIV Tat protein (Xiao et al. 2000), which binds TIP30 (Xiao et al. 1998). TIP30 can associate with an RNAPII-SRB complex and may be a gene-specific co-factor because it is also required for up-regulation of a subset of apoptotic genes but not for basal transcription (Xiao et al. 2000).

3.1.8 CDC2

CDC2 was the first CTD kinase to be described (Cisek and Corden 1989) and phosphorylates the RNAPII CTD on Ser2 and Ser5 residues in vitro (Zhang and Corden 1991). CDC2 preferentially phosphorylates the less conserved carboxy-terminal half of the CTD in vitro and has very limited activity towards the amino-terminal consensus repeats of the CTD (Rickert et al. 1999). Replacing Ser7 with a basic residue such as Lys in a CTD-4 peptide enhances the serine kinase activity of CDC2 by about 20-fold (Trigon et al. 1998).

Phosphorylation of yeast RNAPII by CDC2 can inhibit transcription in vitro (Gebara et al. 1997). CDC2 forms a complex with cyclinB, known as MPF (mitosis promoting factor) that triggers entry into mitosis and CTD phosphorylation. Therefore, phosphorylation of the CTD by CDC2 might be involved in cell cycle regulation (see below).

3.1.9 c-Abl and Abl-related gene product

The proto-oncogene c-Abl can phosphorylate approximately 30 Tyr1 residues of the RNAPII CTD heptapeptide repeat YSPTPS in vitro (Baskaran et al. 1993).
Moreover, transient overexpression of c-Abl results in an increase in the phospho-Tyr content of RNAPII, and c-Abl can associate with the large subunit of RNAPII (Baskaran et al. 1996). The function of c-Abl depends on its tyrosine kinase domain, its SH2 domain and a CTD-interacting domain at the carboxy-terminus of c-Abl (Baskaran et al. 1996). Mammalian c-Abl can only phosphorylate the CTD of human RNAPII and is not active on yeast or Drosophila RNAPII (Baskaran et al. 1999). This species specificity is determined by extreme carboxy-terminal CTD repeats that are not conserved through evolution (Baskaran et al. 1999). It has been suggested that Tyr phosphorylation of the CTD by c-Abl might play a role in transcription initiation (Baskaran et al. 1996) or elongation (Baskaran et al. 1999).

Tyrosine phosphorylation of RNAPII is still observed in a c-Abl negative cell line derived from the embryo of a c-Abl null mutant mouse, suggesting that there is another physiological CTD-tyrosine kinase. Consistent with this, the Abl-related gene product (Arg) can also phosphorylate a GST-CTD fusion protein in vivo and in vitro although effects on phosphorylation of the RNAPII CTD have not been demonstrated (Baskaran et al. 1997a). Phosphorylation of the CTD by c-Abl occurs in response to genotoxic agents (see below).

3.2 CTD phosphatase

In spite of the major focus on CTD kinases and their influence on gene expression, relatively little is known concerning the activity and regulation of CTD phosphatases. The main objective of my thesis work was the molecular characterization of a CTD
phosphatase from *Saccharomyces cerevisiae* and therefore I shall discuss the available experimental data dealing with this enzyme in some detail.

CTD phosphatase was first purified from HeLa cell extracts by assaying for activity on RNAPII phosphorylated with a partially purified serine/threonine CTD kinase (Chambers and Dahmus 1994). The assay employed a mobility shift on RNAPIIO that was further modified by a $^{32}$P-label on the most carboxy-terminal Ser of its largest subunit (Chambers and Dahmus 1994). This Ser residue is flanked by acidic residues and is an ideal substrate for CK-II. Phosphorylation at this site does not seem to alter the activity of RNAPII and therefore is a convenient tool for tagging the enzyme. Using this assay, a CTD phosphatase was partially purified that had an apparent molecular weight of 200000 Da and required magnesium ions for activity (Chambers and Dahmus 1994). The enzyme was also sensitive to okadaic acid and hence classified as a type 2C phosphatase (Chambers and Dahmus 1994). This CTD phosphatase is specific for dephosphorylating Ser and Thr residues within the consensus repeats of the CTD, as it does not dephosphorylate the Ser residue phosphorylated by CK-II (Chambers and Dahmus 1994). CTD phosphatase can dephosphorylate RNAPIIO purified from calf thymus or RNAPIIO prepared from RNAPIIA *in vitro* by phosphorylation with two different HeLa cell CTD kinases whose exact nature is uncertain (Chambers and Dahmus 1994). RNAPII contains a docking site for CTD phosphatase that is essential for CTD phosphatase activity and distinct from the CTD (Chambers et al. 1995). This conclusion is supported by the following observations: phosphorylated recombinant CTD is not a substrate for CTD phosphatase; RNAPIIB (which lacks the CTD) and RNAPIIA are
competitive inhibitors of CTD phosphatase activity and RNAPII can form a stable complex with CTD phosphatase (Chambers et al. 1995).

Two general transcription factors regulate CTD phosphatase activity. TFIIF stimulates CTD phosphatase activity five-fold (Chambers et al. 1995). This stimulatory activity is derived from the carboxy-terminal residues 358-517 of the RAP74 subunit of human TFIIF (Chambers et al. 1995). TFIIB prevents stimulation by TFIIF but does not have an effect on CTD phosphatase in the absence of TFIIF (Chambers et al. 1995). The interaction of TFIIB and RAP74 with CTD phosphatase will be also described in detail in Chapter 3 of this thesis.

The studies described above were performed using a preparation of partially purified human CTD phosphatase containing a number of polypeptides with unknown identities. Subsequently, using RNAPII O phosphorylated by the general transcription factor TFIIF as a substrate for assays, a CTD phosphatase was purified from extracts of S. cerevisiae (Chambers and Kane 1996). Two fractions essential for CTD phosphatase activity were resolved: one contained TFIIF but could not be replaced with highly purified yeast TFIIF (Chambers and Kane 1996); the other essential fraction contained a polypeptide with an apparent mass of 100/103 kDa that co-purified with the activity (Chambers and Kane 1996). This polypeptide was similar in amino acid sequence for most of its length (32% identity and 54% similarity) to a human protein cloned on the basis of its interaction with the RAP74 subunit of human TFIIF (Archambault et al. 1997). The binding site in RAP74 for this human protein was contained within the minimal portion of RAP74 needed for RAP74 to stimulate dephosphorylation of
RNAPIIIO by human CTD phosphatase (Chambers et al. 1995; Archambault et al. 1998b). The homologous yeast protein also interacts with TFIIF (Archambault et al. 1997) and a detailed investigation of this interaction will be presented in Chapter 3 of this thesis. These new proteins were given the name FCP1, an acronym for TFIIF-associating component of CTD phosphatase (Archambault et al. 1997; Archambault et al. 1998b).

Although the requirement for divalent cations and the resistance to okadaic acid and vanadate classified this protein phosphatase as a type 2C, neither the sequence of yeast Fcp1p (Archambault et al. 1997) nor that of human FCP1a (Archambault et al. 1998b) resembled that of any known protein phosphatase of the PP2C or another family (Barford et al. 1998). Therefore, it was not known whether Fcp1p was a phosphatase enzyme or an essential regulatory component of the enzyme. Human FCP1 co-purified with the human CTD phosphatase activity described above through several column purification steps and through subsequent affinity purification to virtual homogeneity on a GST-RAP74 column (Archambault et al. 1998b). Removal of a large portion of the human FCP1 from HeLa nuclear extracts by immunodepletion using anti-FCP1 antibodies reduced the CTD phosphatase activity of these extracts by 3-4 fold (Archambault et al. 1998b).

All these studies therefore supported the idea that Fcp1p is an essential subunit of CTD phosphatase and could be the catalytic subunit of a CTD phosphatase complex. The experiments that I describe in Chapter 2 of this thesis show that Fcp1p is indeed a CTD phosphatase and the founding member of a new family of eukaryotic protein phosphatases. Independent experiments with the human system showed that FCP1 is the
human CTD phosphatase, although residues at the amino-terminal region not present in the original FCP1a clone (Archambault et al. 1998b) are necessary for the catalytic activity (Cho et al. 1999).

Recent studies have focused on the regulation of CTD phosphatase activity during the transcription cycle. A role for CTD phosphatase in recycling RNAPII has been suggested based on in vitro transcription reactions (Cho et al. 1999). FCP1 dephosphorylates RNAPII in vitro to allow efficient incorporation of RNAPII into transcription initiation complexes and this leads to increased transcription (Cho et al. 1999). I will describe below evidence for a possible involvement of FCP1 in transcription elongation.

3.3 Antibodies used to study CTD phosphorylation

Investigation of CTD phosphorylation, especially in vivo, but also in vitro, relies heavily on the differential binding of various antibodies to various epitopes present on the CTD. The specificity of the most commonly used antibodies has been determined by binding to GST-CTD fusion proteins with wild-type CTD repeats or various Ser to Ala or Ser to Glu substitutions. It should be noted that the fusion protein substrate was phosphorylated with CDC2 kinase (Patturajan et al. 1998a), and contained only 15-16 repeat sequences.

Monoclonal antibody 8WG16 primarily recognizes the unphosphorylated CTD, and Ser2 seems to be an important feature of the 8WG16 epitope (Patturajan et al. 1998a). Phosphorylation of the CTD inhibits its interaction with 8WG16. H5 and H14
are phospho-specific monoclonal antibodies that recognize phospho-Ser2 (H5) and phospho-Ser5 (H14), respectively (Patturajan et al. 1998a). The monoclonal antibodies B3 and MARA3 recognize the phosphorylated CTD when Ser2 or Ser5 or both are phosphorylated, although neither phosphate is essential (Patturajan et al. 1998a). Monoclonal antibody CC3 is similar to H5, but differs in some respects: phosphorylation of Ser2 along with other, uncharacterized determinants is important (Patturajan et al. 1998a), and CC3 seems to preferentially bind the non-consensus repeats at the carboxy-terminus of the CTD (Dubois et al. 1997).

3.4 Complexes involved in initiation of transcription

3.4.1 The Srb/mediator complex, RNAPII holoenzymes, and the CTD

The identification of a pre-assembled RNAPII transcription complex that integrates the signals from gene-specific activators and repressors is an intriguing example of how genetics and biochemistry can converge. Using a genetic approach, mutations in the SRB (suppressor of RNA polymerase B) genes were identified as suppressors of growth defects associated with truncations of the yeast RNAPII CTD (Nonet and Young 1989). Dominant suppressors were identified in the SRB2, SRB4, SRB5 and SRB6 genes (Carlson 1997). The analysis of recessive mutations yielded additional alleles of srb4 and srb6 and also identified SRB7, SRB8, SRB9, SRB10 and SRB11. SRB4, SRB6 and SRB7 are essential for viability. SRB2 and SRB5 are non-essential genes, although deletions cause various phenotypes including slow growth and cold-sensitivity (Carlson 1997). Only dominant alleles of SRB2 and SRB5 have been isolated as suppressors of the CTD.
truncation. $SRB8$, $SRB9$, $SRB10$ and $SRB11$ are also non-essential genes, and deletions of these genes cause slow growth, flocculence and mild temperature- and cold-sensitive phenotypes (Carlson 1997). Only recessive alleles were recovered as suppressors of the CTD truncation. $SRB10$ and $SRB11$ encode yeast homologs of mammalian cyclin-dependent kinase 8 (CDK8) and cyclinC and function as a CTD kinase (Liao et al. 1995) (described above). The genes encoding Srb proteins as well as genes encoding components of the mediator and other RNAPII associated complexes are listed in Table 3. The table is adapted from (Myers and Kornberg 2000) and (Carlson 1997) with additional information from YPD (www.proteome.com). Major abbreviations used are yMed (yeast mediator according to the Kornberg laboratory), yHolo (yeast RNAPII holoenzyme according to the Young laboratory) and yPAF (yeast Paf1p containing complex according to the Jaehning laboratory).
<table>
<thead>
<tr>
<th>Gene</th>
<th>Aliases</th>
<th>Essential for viability</th>
<th>Size (kDa)</th>
<th>Complex</th>
<th>Functional Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGR1</td>
<td></td>
<td>Yes</td>
<td>123</td>
<td>yMed</td>
<td>Required for glucose repression; carboxy-terminus anchors Mediator sub-complex containing Sin4p, Gal1p, Pgd1p, Med2p</td>
</tr>
<tr>
<td>GAL1</td>
<td>SPT13</td>
<td>No</td>
<td>120</td>
<td>yMed, yPAF</td>
<td>Dual role in repression and activation; might directly interact with activation domains</td>
</tr>
<tr>
<td>SIN4</td>
<td>SSN4, TSF3,</td>
<td>No</td>
<td>111</td>
<td>yMed</td>
<td>Dual role in repression and activation</td>
</tr>
<tr>
<td></td>
<td>SD13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRB4</td>
<td></td>
<td>Yes</td>
<td>78</td>
<td>yMed</td>
<td>Global role in RNAPII transcription; dominant and recessive suppressor of CTD truncations; interacts with MED6; target for Gal4p</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>yHolo</td>
<td></td>
</tr>
<tr>
<td>MED1</td>
<td></td>
<td>No</td>
<td>64</td>
<td>yMed</td>
<td>Dual role in repression and activation; possible functional link to Srb10p/Srb11p</td>
</tr>
<tr>
<td>Gene</td>
<td>Mediator Complex</td>
<td>Condensation</td>
<td>Median</td>
<td>yMed</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>------------------</td>
<td>--------------</td>
<td>--------</td>
<td>------</td>
<td>-------------</td>
</tr>
<tr>
<td>MED2</td>
<td></td>
<td>No</td>
<td>48</td>
<td>yMed</td>
<td>Required for specific activators; interaction with mediator dependent upon Pgd1p</td>
</tr>
<tr>
<td>PGD1</td>
<td>MED3, HRS1</td>
<td>No</td>
<td>47</td>
<td>yMed</td>
<td>Suppressor of hyper-recombination; interacts with Med2p, positively acting regulator of transcription; direct target of transcriptional repression by Ssn6p/Tup1p</td>
</tr>
<tr>
<td>SRB5</td>
<td></td>
<td>No</td>
<td>35</td>
<td>yMed, yHolo</td>
<td>Required for efficient transcription initiation (16% of yeast genes); dominant suppressor of CTD truncations; required \textit{in vivo} for transcriptional activation by Gal4p</td>
</tr>
<tr>
<td>MED6</td>
<td>Yes</td>
<td>32</td>
<td>yMed</td>
<td></td>
<td>Interacts with Srb4p and shows specific defects in activation (10% of all yeast genes); required for VP16-activated transcription \textit{in vitro}</td>
</tr>
<tr>
<td>MED7</td>
<td>Yes</td>
<td>32</td>
<td>yMed</td>
<td></td>
<td>Only mediator subunit found in all mediator complexes from higher cells</td>
</tr>
<tr>
<td>MED4</td>
<td></td>
<td>Yes</td>
<td>32</td>
<td>yMed</td>
<td></td>
</tr>
<tr>
<td>ROX3</td>
<td>SSN7, ARE3</td>
<td>Yes</td>
<td>25</td>
<td>yMed</td>
<td>Involved in glucose repression and regulation of stress response; repressor of \textit{HO} transcription</td>
</tr>
<tr>
<td><strong>MED8</strong></td>
<td>Yes</td>
<td>25</td>
<td>yMed</td>
<td>Possibly binds to UAS of <em>SUC2</em> and <em>HXK2</em> genes</td>
<td></td>
</tr>
<tr>
<td><strong>SRB2</strong></td>
<td>No</td>
<td>23</td>
<td>yMed, yHolo</td>
<td>Interacts with Srb4p and Srb5p; dominant suppressor of CTD truncations</td>
<td></td>
</tr>
<tr>
<td><strong>NUT2</strong></td>
<td>Yes</td>
<td>18</td>
<td>yMed</td>
<td>Involved in repression and activation</td>
<td></td>
</tr>
<tr>
<td><strong>MED10</strong></td>
<td>Yes</td>
<td>18</td>
<td>yMed</td>
<td>Involved in repression and activation</td>
<td></td>
</tr>
<tr>
<td><strong>CSE2</strong></td>
<td>No</td>
<td>17</td>
<td>yMed</td>
<td>Involved in stimulation of basal transcription</td>
<td></td>
</tr>
<tr>
<td><strong>SRB7</strong></td>
<td>Yes</td>
<td>16</td>
<td>yMed, yHolo</td>
<td>Recessive suppressor of CTD truncation</td>
<td></td>
</tr>
<tr>
<td><strong>MED11</strong></td>
<td>Yes</td>
<td>15</td>
<td>yMed</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SRB6</strong></td>
<td>Yes</td>
<td>14</td>
<td>yMed, yHolo</td>
<td>Binds Srb4p; generally required for RNAPII transcription; dominant and recessive suppressor of CTD truncations</td>
<td></td>
</tr>
<tr>
<td><strong>SRB8</strong></td>
<td>SSN5, ARE2</td>
<td>No</td>
<td>167</td>
<td>yHolo</td>
<td>Important for negative regulation; recessive suppressor of CTD truncations</td>
</tr>
<tr>
<td><strong>SRB9</strong></td>
<td>SSN2, UME2, SCA1</td>
<td>No</td>
<td>160</td>
<td>yHolo</td>
<td>Important for negative regulation; recessive suppressor of CTD truncations; genetic interaction with Ser2 of CTD</td>
</tr>
</tbody>
</table>
| **SRB10** | SSN3, UME5, ARE1 | No | 63 | yHolo | Important for negative regulation; important for transcriptional activation by Gal4p; recessive suppressor of CTD truncations,
<table>
<thead>
<tr>
<th>Gene</th>
<th>Protein</th>
<th>No/Yes</th>
<th>Cyclin-dependence</th>
<th>Function Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRB11</td>
<td>SSN8, UME3</td>
<td>No</td>
<td>38</td>
<td>yHolo</td>
</tr>
<tr>
<td>CCR4</td>
<td>FUN27, NUT21</td>
<td>No</td>
<td>95</td>
<td>yPAF NOT</td>
</tr>
<tr>
<td>HPR1</td>
<td>TRF1</td>
<td>No</td>
<td>88</td>
<td>yPAF</td>
</tr>
<tr>
<td>PAF1</td>
<td></td>
<td>No</td>
<td>52</td>
<td>yPAF</td>
</tr>
<tr>
<td>CDC73</td>
<td></td>
<td>No</td>
<td>44</td>
<td>yPAF</td>
</tr>
<tr>
<td>SOH1</td>
<td></td>
<td>No</td>
<td></td>
<td>Human PC2 and SMCC</td>
</tr>
</tbody>
</table>
The nine Srb proteins can exist in a complex with RNAPII (Koleske and Young 1994). This RNAPII holoenzyme complex also contains a subset of the general transcription factors (TFIIB, TFIIF, and TFIIH), Gal1lp, and probably other polypeptides and can support basal transcription when supplemented with TBP and TFIIE (Koleske and Young 1994). Importantly, despite the lack of TBP-associated proteins (TAFIIS), this RNAPII holoenzyme can bind the VP16 activation domain (Hengartner et al. 1995) and support activated transcription in response to Gal4-VP16 in vitro (Koleske and Young 1994). Some preparations of this Srb complex also contain components of the SWI/SNF complex involved in chromatin remodeling (Wilson et al. 1996), although this association may be transient and not very stable (Cairns et al. 1996b).

A search for factors in yeast extracts that confer responsiveness to activator proteins led to the discovery of mediator and provided independent evidence for the existence of preassembled RNAPII complexes. Mediator was purified as part of a mediator–RNAPII complex that also contained the general transcription factor TFIIF but no other general transcription factors (Kim et al. 1994). Upon addition of the remaining general transcription factors, this holoenzyme also supports basal and activated transcription in vitro (Kim et al. 1994).

Interaction of mediator with RNAPII is partly mediated by the CTD (Myers et al. 1998). This association can be disrupted by the 8WG16 monoclonal antibody that primarily recognizes the hypophosphorylated form of the CTD (Kim et al. 1994; Svejstrup et al. 1997). Mediator can directly interact with a GST-CTD fusion protein in vitro (Myers et al. 1998), but additional multiple contact points between RNAPII and
mediator exist that are independent of the CTD (Asturias et al. 1999). Electron microscopy showed that mediator forms a compact, ellipsoidal structure, with dimensions of about 400 by 200\(\text{Å}\) (Asturias et al. 1999). In the presence of RNAPII, mediator seems to unfold to a crescent of density that partially surrounds the RNAPII (Asturias et al. 1999).

Three functional activities have been associated with mediator. Mediator not only makes activated transcription possible but also stimulates basal transcription about tenfold and stimulates CTD phosphorylation by the TFIH kinase by thirtyfold to fiftyfold (Kim et al. 1994). The RNAPII CTD is required for mediator to stimulate basal and activated transcription (Myers et al. 1998). Consistent with this, the RNAPII found associated with Srb/mediator complexes contains a hypophosphorylated CTD (Kim et al. 1994; Koleske and Young 1994). In contrast, mediator is not associated with elongating RNAPII that contains a hyperphosphorylated CTD (Svejstrup et al. 1997) (see below).

Differences in protein composition between the Srb complex and the mediator may be simply a consequence of different purification strategies. Mediator lacks Srb8p, Srb9p, Srb10p, and Srb11p, but contains Srb2p, Srb4p, Srb5, Srb6p, Srb7p and nine other polypeptides previously identified as either negative or positive regulators of transcription, including Gal11p, Rgr1p, Sin4p, and Pgd1p. Seven additional mediator subunits are encoded by MED genes which were not identified previously through genetic analysis (Myers and Kornberg 2000). Mediator can be dissociated into two tightly associated sub-complexes. The Srb module consists of Med6p, Rox3p, Srb2p, Srb4p, Srb5p and Srb6p (Lee and Kim 1998). The Srb proteins in this sub-complex are all
products of genes that were identified as dominant gain-of-function suppressors of the CTD truncation in the original genetic screen. The Rgrlp-associated module consists of Rgrlp, Srb7p, Medlp, Med2p, Med4p, Med7p, Med8p, Med11p, Cse2p, Nut2p, and the closely associated Gnl1p-Sin4p-Pgd1p sub-complex (Lee and Kim 1998). Figure 1 illustrates the modular structures of the yeast Srb/mediator and Paf1p (described below) complexes.
Figure 1. Protein complexes involved in transcription initiation in *Saccharomyces cerevisiae*. Yeast Srb/mediator has various modules that have been identified through genetic and biochemical studies and these are represented in different colors. The functions of the Srb/mediator modules are mediated partly through the CTD, whereas the Paf1p complex interacts with RNAPII independently of the CTD. Examples of upstream regulatory proteins and signaling pathways that function through these complexes are indicated. Straight arrows indicate direct physical and/or functional interactions. The general transcription factors are shown in association with RNAPII. See text for further details.
3.4.2. The Srb/mediator complex and transcriptional regulation

I will next summarize genetic and biochemical evidence for Srb/mediator function in transcription. The RNAPII holoenzyme appears to have a global role in transcription. Mutations in *srb4* and *srb6* result in a rapid decline of poly(A)^+ mRNA levels upon shift of the cells to the non-permissive temperature (Thompson and Young 1995), thus suggesting that most, if not all, transcription initiation *in vivo* involves RNAPII holoenzyme. Consistent with this, the gene expression defects in an *srb4* strain are nearly indistinguishable from those in a strain with a mutation in the largest subunit of RNAPII (Holstege et al. 1998). The *CUP1, HSP82* and *SSA4* genes are exceptions to this rule of Srb4p-dependence (Lee and Lis 1998; McNeil et al. 1998). However, even these genes require some components of mediator, such as Rgr1p, for transcriptional activation (Lee et al. 1999a). In contrast to the nearly general role of Srb4p and Srb6p, Srb5p, contained within the same module of mediator, regulates approximately 16% of all yeast genes under normal growth conditions (Holstege et al. 1998). Similarly, 10% of all genes are dependent on Med6p (Myers et al. 1999). However, 90% of the genes that have reduced expression levels in either the *srb5Δ* or the *med6* strains are different from each other, suggesting that these factors regulate largely different subset of genes *in vivo* (Myers and Kornberg 2000).

The (Med6p, Rox3p, Srb2p, Srb4p, Srb5p, Srb6p) module may also be a common target for upstream activator proteins. Med6p is required for transcriptional activation by various activators *in vivo* (Han et al. 1999), and RNAPII holoenzymes from a *med6* strain
have defects in VP16-activated transcription in vitro (Lee et al. 1997). Srb4p directly interacts with the activation domain of yeast Gal4p (Koh et al. 1998), and mutations in the Gal4p-binding domain of Srb4p can partly restore Gal4p-binding and activation in vivo by a Gal4p derivative bearing a mutant activation (Koh et al. 1998). Thus, taken together, this data indicate an important role for the Srb/mediator in response to transcriptional activators bound to upstream regions. Consistent with this, the mediator dependence of activated transcription seems to reside in the upstream regulatory elements and not in the core promoter sequences. Srb2p functions through the upstream activating sequence (UAS) of the INO1 promoter (Scafe et al. 1990) and Med2p functions through a GAL UAS (Myers et al. 1999).

Less is known about the (Rgrlp, Srb7p, Med1p, Med2p, Med4p, Med7p, Med8p, Med11p, Cse2p, Nut2p) module and the associated (Pgd1p, Sin4p and Gal11p) complex. Although some of these proteins are encoded by essential genes, they have not yet been shown to be generally required for transcription. Instead, most of them have been implicated in negative and/or positive regulation at specific promoters in vivo. Rgrlp is necessary for repression by α2 and for negative regulation of meiosis genes, the HO gene, and genes lacking UAS, but is also involved in activation of certain genes (Carlson 1997). The carboxy-terminus of Rgrlp anchors the Gal11p-Sin4p-Pgd1p complex to the mediator (Li et al. 1995; Myers et al. 1999). SRB7 has not yet been extensively characterized in yeast, but its mouse homolog is expressed in all tissues and is essential for cell viability and for murine embryonic development (Tudor et al. 1999). Med1p is
important for \textit{SNF1}-dependent glucose repression, but is also necessary for normal levels of \textit{GAL} gene activation (Balciunas et al. 1999). Med1p might be regulated by Srb10p/Srb11p because a LexA-DNA binding domain fusion to Med1p can only stimulate high levels of transcription in an \textit{srb11Δ} strain (Balciunas et al. 1999). Med7p is the only yeast mediator subunit whose homolog has been identified in all mammalian mediator complexes (Malik and Roeder 2000) (see below). In yeast, conditional depletion of Med7p results in complete arrest of cell division (Belli et al. 1998), a unique phenotype among mediator mutants. It has been suggested that Med8p can bind directly to the upstream regulatory elements of the \textit{SUC2} and \textit{HXX2} genes (Chaves et al. 1999; Moreno-Herrero et al. 1999). Med11p is specifically required for \textit{MFalpha1} transcription (Han et al. 1999). A yeast strain lacking Cse2p has a defect in Bas1p/Bas2p-mediated transcription of amino acid biosynthetic genes and mediator prepared from this strain has a slight defect in the stimulation of basal transcription (Han et al. 1999). Nut2p also has a putative dual role in activation and repression (Myers and Kornberg 2000). A strain with a deletion of \textit{med2} has a greater than twofold defect in the expression of about 4% of all yeast genes (Myers et al. 1999), as well as a defect in \textit{GAL} gene activation \textit{in vivo} (Myers et al. 1999). Mediator lacking Med2p cannot support Gal4-VP16 activation \textit{in vitro} but is able to support activation by Gcn4p (Myers et al. 1999). Med2p probably associates with the mediator through direct interactions with Pgd1p, Sin4p, and Gal11p (Lee et al. 1999b; Myers et al. 1999).

Together with Med2p, the Sin4p-Pgd1p-Gal11p module may constitute another
activator-binding entity in the mediator complex. This module appears to be dispensable for basal transcription but is required for selective effects on activated transcription. For example, activation by both VP16 and Gcn4p is not supported by mediator purified from a sin4Δ strain (which also lacks Med2p and Pgd1p). HIS gene expression is regulated by Gcn4p and is defective in a sin4Δ strain, consistent with the in vitro transcription results. 

GAL11 has been identified in a number of genetic screens for positive and negative regulators of transcription (Carlson 1997) and may interact with acidic activation domains in vitro, consistent with a role in GAL gene activation in vitro (Lee et al. 1999b).

Consistent with an additional role for mediator in the negative regulation of transcription, the Pgd1p subunit of the Sin4p-Pgd1p-Gal11p module can also directly interact with the yeast corepressor Ssn6p/Tup1p, and this contact may prevent recruitment of the holoenzyme to the core promoter (Papamichos-Chronakis et al. 2000). Ssn6p/Tup1p is targeted to specific promoters by interacting with site-specific DNA-binding proteins and several genetic screens have identified genes whose products affect repression by Ssn6p/Tup1p in vivo and are subunits of the Srb/mediator complex (Smith and Johnson 2000).

Although the yeast Srb and mediator complexes are remarkably similar and partly overlapping, there are some notable differences. The Srb8-11 proteins are absent from the mediator (Myers et al. 1998). The Srb8-11 genes were identified in several genetic screens as negative regulators of transcription, and loss-of-function mutations in these genes cause derepression of certain promoters (Carlson 1997). Although certain mutations
in *rgrl*, *sin4*, *rox3*, and *gal11*, which encode proteins found in different modules of mediator, lead to similar defects in transcriptional repression, these two groups of genes and their products are clearly distinct: no alleles of *RGR1*, *SIN4*, *ROX3* and *GAL11* have been recovered as suppressors of CTD truncations, and only mutated alleles of *rgrl*, *sin4*, *rox3* and *gal11* derepress promoters lacking a UAS (Carlson 1997). Mutated alleles of *srb8-11* cannot derepress this particular class of promoter.

The Srb8-11p complex may have a negative role because the Srb10p/Srb11p CDK/cyclin pair phosphorylates the CTD. Srb10p is uniquely able to phosphorylate the CTD in purified RNAPII holoenzymes prior to template binding, and this phosphorylation inhibits subsequent transcription by the RNAPII holoenzyme (Hengartner et al. 1998). Pre-incubation with ATP of RNAPII holoenzyme purified from a *SRB10* strain leads to inhibition of transcription, whereas pre-incubation of a RNAPII holoenzyme from a *srb10* mutant strain does not (Hengartner et al. 1998). Srb10p does not appear to inhibit transcription after the formation of a stable PIC (Hengartner et al. 1998). Consistent with a function of Srb10p/Srb11p in the negative regulation of transcription, fusion to a heterologous DNA-binding domain and recruitment to a promoter of an Srb10p with an inactive kinase due to a mutation in its ATP-binding site results in a much larger stimulation of transcription than recruitment of wild-type Srb10p (Hengartner et al. 1998). Also consistent with this negative role for Srb10p, data presented in Chapter 4 of this thesis suggest that Srb10p and the CTD phosphatase Fcp1p have opposing activities in the regulation of the phosphorylation state of free RNAPII CTD.
Differential phosphorylation of the CTD may be involved in the regulation of transcription by the Srb/mediator complex. \textit{SRB9} was identified as a suppressor of mutations in which Ser2 was replaced with Ala or Glu in multiple CTD repeats (Yuryev and Corden 1996). The suppression was specific for Ser2 substitutions because Ser5 substitutions were not suppressed (Yuryev and Corden 1996). Dominant alleles of \textit{SRB2}, \textit{SRB4}, \textit{SRB5} and \textit{SRB6} also cause the same suppression (Yuryev and Corden 1996). It is also possible that the CTD, through its association with the Srb/mediator complex, overcomes general negative regulators of transcription. First, Srb2p and Srb5p are required for basal transcription in a crude extract (Thompson et al. 1993), even though basal transcription in a highly purified system can occur without mediator (Sayre et al. 1992; Svejstrup et al. 1994). Second, as mentioned above, the RNAPII CTD is similarly not required for basal transcription in a reconstituted system but is needed in a crude system. Consistent with this idea, an \textit{srb4} mutation can be suppressed by recessive mutations in \textit{bur6}, \textit{ncb2}, \textit{not1}, \textit{not3}, \textit{not5} and \textit{cafl} (Lee et al. 1998). These genes encode subunits of NC2 (Gadbois et al. 1997; Prelich 1997) and the NOT complex (Collart and Struhl 1994), both of which are general negative regulators that interact with TBP.

Mediator purified from the fission yeast \textit{Schizosaccharomyces pombe} is smaller than the \textit{S. cerevisiae} mediator but contains the \textit{S.pombe} homologs of Rgr1p, Srb4p, Med7p and Nut2p (Spahr et al. 2000). As was the case for \textit{S.cerevisiae}, the presence of mediator in a \textit{S.pombe} RNAPII holoenzyme can stimulate phosphorylation of the CTD by TFIIH (Spahr et al. 2000).
3.4.3. Other RNAPII complexes

Recently, a distinctly different form of RNAPII complex was purified from yeast cells by binding to RNAPII immobilized on a matrix containing the 8WG16 antibody that recognizes hypophosphorylated CTD (Shi et al. 1996; Wade et al. 1996; Shi et al. 1997). This complex lacked most Srb/mediator proteins, which may be displaced by the antibody (Wade et al. 1996). The complex contained the general transcription factors TFIIB and TFIIF, the mediator component Gal1lp and a novel Paf1p sub-complex consisting of Paf1p, Cdc73p, Hpr1p and Ccr4p (Shi et al. 1997; Chang et al. 1999). The genes encoding subunits of this complex have been included in Table 3 and a diagram of the complex in Figure 1, both presented above.

Although they differ in composition, several genetic connections between the Paf1p complex and the Srb/mediator complex exist. Deletions of paf1 and ccr4 cause lethality in yeast when combined with a deletion of srb5 (Chang et al. 1999), suggesting that the Srb/mediator and Paf1p complexes have partly overlapping functions. Ccr4p is involved in the negative and positive regulation of many genes and also associates with the NOT complex (Bai et al. 1999). Therefore, CCR4 is also genetically connected to mediator through its association with the products of the NOT genes that had been shown to suppress an srb4 mutation (Lee et al. 1998). Deletion of hprl leads to a hyperrecombination phenotype as well as to a defect in transcriptional elongation in a lacZ reporter gene (Chavez and Aguilera 1997). The hyper-recombination phenotype can be suppressed by mutant alleles of pgd1, therefore providing an additional genetic link between this Paf1p complex and the mediator (Santos-Rosa and Aguilera 1995).
Additional genetic analysis of the Paf1p complex has suggested that it is a downstream target of the protein kinase C signaling cascade (Chang et al. 1999). The Hpr1p component of this complex also provides a link to the human co-activator complex SMCC through the human homolog of Soh1p (described below). Mutations in the sohl gene can suppress the hyper-recombination phenotype of a hpr1Δ strain (Fan et al. 1996).

3.4.4 Mammalian RNAPII holoenzyme, co-activator complexes and the RNAPII CTD

Equivalents of the yeast RNAPII holoenzyme have been purified from both rat and human cells. The first mammalian RNAPII holoenzyme was purified by immunoaffinity chromatography from rat liver nuclear extracts. This holoenzyme contained all of the general transcription factors and the RNAPIIA and could accurately initiate transcription in vitro (Ossipow et al. 1995). A more complex form of the human RNAPII holoenzyme was purified on the basis of its interaction with the elongation factors TFIIS and Elongin A (Pan et al. 1997). All of the general transcription factors were present in this holoenzyme along with RNAPIIA, human homologs of some of the Srb proteins and the cleavage and polyadenylation factors CPSF and CstF (McCracken et al. 1997b; Pan et al. 1997). This complex could support activated transcription by Gal4-VP16 and Gal4-Sp1 in vitro. Other mammalian RNAPII complexes have been purified. An emerging theme is heterogeneity in the composition of these complexes, which contain
different subsets of GTFs, Srb/mediator proteins and factors that affect chromatin structure and nucleic acid metabolism (Myer and Young 1998).

Transcriptional co-activators are involved in transcriptional activation by sequence-specific DNA-binding activators but are not necessarily involved in basal transcription. Recently, distinct co-activator complexes have been isolated that did not co-purify with RNAPII holoenzymes and nevertheless provided evidence for the evolutionary conservation of mediator (Hampsey and Reinberg 1999; Malik and Roeder 2000; Myers and Kornberg 2000). Mammalian TRAP was purified to homogeneity on the basis of its ligand-dependent interaction with the thyroid hormone receptor (Fondell et al. 1996; Fondell et al. 1999). The purification of mammalian homologs of yeast Med7p, Srb7p, Srb10p and Srb11p yielded a number of related complexes, including SMCC, which appears identical to TRAP (Sun et al. 1998; Gu et al. 1999; Ito et al. 1999). In all, nine mediator complexes have been purified from mouse and human cells (Malik and Roeder 2000; Myers and Kornberg 2000) (see Table 4). All of the mammalian mediator complexes contain homologs of Med7p and most have homologs of Med6p, Rgr1p, Srb7p and Nut2p. Only the TRAP/SMCC, SUR2 and NAT complexes contain homologs of Srb10p/Srb11p (Sun et al. 1998; Boyer et al. 1999; Gu et al. 1999; Ito et al. 1999). TRAP/SMMC and PC2 also contain a human homolog of the yeast SOHL gene product (Gu et al. 1999). SOHL interacts genetically with RBP2 (Rpb2p), SUA7 (TFIIB) and HPRI (Fan et al. 1996), which encodes a component of the Paflp-RNAPII complex (Chang et al. 1999). A protein with homology to the elongation factor TFIIS is found in some co-activator complexes. I present a summary of prominent and common subunits of
the metazoan co-activator complexes and their relationship to yeast Srb/mediator in Table 4. The information is partly adapted from recent reviews (Malik and Roeder 2000; Myers and Kornberg 2000).
Table 4. Human co-activator complexes and their relationships to yeast

Srb/mediator

<table>
<thead>
<tr>
<th>Factor</th>
<th>TRAP SMCC</th>
<th>PC2</th>
<th>DRIP</th>
<th>ARC</th>
<th>CRSP</th>
<th>NAT</th>
<th>SUR2</th>
<th>Murine MED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rgr1</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Med4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Med6</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Med7</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Nut2</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Srb7</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Srb10/CDK8</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Srb11/CycC</td>
<td>+</td>
<td>-</td>
<td></td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Sohl</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IIS related</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUR2</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Other Subunits</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Consistent with their functions as co-activators of transcription, several subunits of the murine mediator complexes can interact with transcriptional activator proteins, often in a function-dependent manner. These include TRAP220 whose interaction with thyroid receptor is ligand-dependent (Yuan et al. 1998), TRAP80, which interacts with the activation domains of VP16 and p53 (Ito et al. 1999), and SUR2, which interacts with E1A (Boyer et al. 1999).

In contrast to yeast mediator, the activating function of SMCC is independent of the RNAPII CTD (Gu et al. 1999) but is synergistic with PC4 under conditions of limiting TFIIH concentration (Gu et al. 1999). When TFIIH is present at normal levels, SMCC can repress activator-dependent transcription mediated by PC4 (Gu et al. 1999). Thus, both positive and negative regulatory functions of SMCC are independent of the CTD but might be mediated by PC4, itself a co-activator protein with multiple functions (Gu et al. 1999).

NAT also can repress or activate transcription depending on the presence of TFIIH (Sun et al. 1998), and therefore might be closely related to SMCC. NAT can physically and functionally associate with RNAPII lacking the CTD, suggesting that the CTD is clearly not the only target of NAT (Sun et al. 1998). However, association of NAT with RNAPII involves the CTD because prior phosphorylation of the CTD precludes an interaction between NAT and RNAPII (Sun et al. 1998). NAT itself can phosphorylate the RNAPII CTD at Ser2 and Ser5 residues, probably because of the presence of CDK8/cyclinC (Sun et al. 1998). Repression of activated transcription by NAT is dependent on CDK8 kinase activity and might be connected to CDK8-dependent
phosphorylation of the cyclinH subunit of TFIH, resulting in an inhibition of the ability of TFIH phosphorylate the CTD \textit{in vitro} and participate in transcription (Akoulitchev et al. 2000).

The murine mediator contains an additional protein related to the Ring-3 protein, a mitogen-activated nuclear protein kinase, raising the possibility that mediator could be an end-point of signal transduction pathways (Jiang et al. 1998). Although murine mediator has fewer proteins than yeast mediator, it forms a similar structure, supporting the evolutionary conservation of mediator (Asturias et al. 1999).

It should be noted that none of these mammalian coactivator/mediator complexes contains homologs of the products of the \textit{SRB2}, \textit{SRB4}, \textit{SRB5} and \textit{SRB6} genes identified as dominant suppressors of yeast CTD truncations. This entire module of the yeast complex either does not exist in mammals or has yet to be discovered. Additionally, whereas the yeast RNAPII holoenzyme/mediator complex can support activated transcription in the presence of TBP (Kim et al. 1994; Koleske and Young 1994), the mammalian co-activator complexes still require TFIID in order to support activated transcription \textit{in vitro} in purified systems.

3.4.5. Recruitment of RNAPII as a mechanism of transcriptional activation

The discovery of the yeast RNAPII holoenzyme led to the proposal that simple recruitment of the transcription complex by upstream DNA-binding activator proteins is the pre-dominant mechanism by which activators work (Ptashne and Gann 1997). This concept has been supported by a number of experiments, although recent results indicate
that transcriptional activation involves other steps. I shall discuss the original concept in some detail because it was used as the basis for the artificial recruitment assays with Fcp1p that are presented in Chapter 3 of this thesis.

Artificially recruiting the TBP or TAF\textsubscript{II}0 subunit of TFIIID to a yeast promoter by fusing it to the DNA-binding domain of either Gal4p or LexA results in a high level of transcription and bypasses the requirement for an activator (Chatterjee and Struhl 1995; Klages and Strubin 1995; Xiao et al. 1995; Gonzalez-Couto et al. 1997). Although the presence of TFIIID in RNAP\textsubscript{II} holoenzyme is controversial, using DNA-binding domain fusion proteins to recruit holoenzyme components such as Gal11p, TFIIB, Sin4p, Srb2p, Srb6p, Srb11p and Rox1p also results in high levels of activation (Jiang and Stillman 1992; Kuchin et al. 1995; Farrell et al. 1996; Song et al. 1996; Gonzalez-Couto et al. 1997). In all cases, simply connecting an activation domain to the holoenzyme component in the absence of a DNA-binding domain does not result in activation (Keaveney and Struhl 1998). Artificial recruitment can also be mediated by a non-covalent interaction between a DNA-tethered peptide and the transcription apparatus. For example, strong activation is achieved when the gal11p mutation in RNAP\textsubscript{II} holoenzyme enables it to interact artificially with a promoter-bound portion of the dimerization domain of Gal4p (Barberis et al. 1995; Farrell et al. 1996). Most of these "non-classical" activators also synergize with classical activation domains when appropriate reporter constructs are used.

There are a definite limitations to the simple recruitment model. It appears that activation by artificial recruitment is highly dependent on the positions of the activator-
binding sites, the promoter sequence, and the coding sequences downstream of the promoter (Gaudreau et al. 1999), whereas strong, natural activators can function in all these situations (Gaudreau et al. 1999). Also, only artificial recruitment of TBP activates transcription in human cells whereas recruitment of various components of RNAPII holoenzyme fails to do so (Nevado et al. 1999; Dorris and Struhl 2000).

Furthermore, physical recruitment of RNAPII holoenzyme is not sufficient for activation. A RNAPII holoenzyme lacking Srb5p still binds to the activation domain of VP16 as strongly as wild-type holoenzyme, yet is unable to support basal and activated transcription (Lee et al. 1999b). This defect cannot be overcome by artificial recruitment to a promoter of RNAPII holoenzyme that lacks Srb5p (Lee et al. 1999b). As well, mediator that lacks Med2p or Pgdlp still contains Srb4p, which is directly targeted by the activation domain of Gal4p (Koh et al. 1998), yet is unable to support activation of the GAL genes in vivo (Myers et al. 1999). These experiments suggest that mediator has important functions after the recruitment of RNAPII holoenzyme.

Using a transcription system based on nuclear extracts and immobilized templates, the formation of the preinitiation complex and the influence of activator proteins was investigated (Ranish et al. 1999). Two stable, intermediate pre-initiation complexes were identified: first, TFIID and TFIIA assembles at the promoter independently of RNAPII holoenzyme; the holoenzyme was then recruited to the promoter, and this was dependent on the CTD, Srb2p, Srb4p, Srb5p and certain portions of TFIIB (Ranish et al. 1999). The Gal4-AH model activator stimulated the rate of PIC formation by enhancing binding to the promoter of TFIIA and TFIID, whereas Gal4-VP16 enhanced the recruitment of
TFIID, TFIIA and the holoenzyme (Ranish et al. 1999). Importantly, transcription extracts from yeast strains with certain mutations in the amino-terminus of TFIIB have a severe defect in transcription but no defect in PIC formation (Cho and Buratowski 1999; Ranish et al. 1999). This important post-assembly function of TFIIB might be also influenced by activators because the same region of TFIIB has also been implicated in the response to activators (Roberts and Green 1994; Wu et al. 1999). In this respect, it is also notable that I shall describe an interaction between TFIIB and Fcp1p in Chapter 3 of this thesis.

In summary, although recruitment of the transcription machinery by activators is important for transcriptional activation, activation likely also involves effects on promoter melting, initiation and promoter clearance. The processivity of elongation by RNAPII might also be regulated by activator proteins. Finally, I shall not discuss chromatin remodeling, even though it is clear that this an important step transcription activation (Allard et al. 1999; Cosma et al. 1999; Ikeda et al. 1999; Natarajan et al. 1999; Neely et al. 1999; Wallberg et al. 1999; Whitehouse et al. 1999; Yudkovsky et al. 1999).

3.5 Transcription Elongation and the CTD

Following transcription initiation, RNAPII must escape the promoter. However, the very early elongation complexes containing transcripts less than approximately 9 nucleotides in length are unstable and prone to abortive initiation and transcriptional arrest (Holstege et al. 1997; Keene and Luse 1999). Interestingly, TFIIIE, the XPB helicase subunit of TFIIH (Dvir et al. 1996; Dvir et al. 1997; Kugel and Goodrich 1998;
Kumar et al. 1998; Moreland et al. 1999; Bradsher et al. 2000; Seroz et al. 2000) and TFIIF (Yan et al. 1999) are required for RNAPII to escape into the processive elongation phase.

Regulation of the processivity of RNAPII elongation complexes might play a role in the expression of many genes. Whereas mRNA synthesis \textit{in vivo} is estimated to proceed at rates of approximately 1000-2000 nucleotides per minute, the rate observed \textit{in vitro} even under optimal reaction conditions is only about 100-300 nucleotides per minute (Izban and Luse 1992). Transcription elongation \textit{in vitro} is also interrupted by frequent pausing, which sometimes ends with arrest when RNAPII slides backward on the DNA template (Uptain et al. 1997). Transcription elongation factors that either enhance the overall rate of mRNA synthesis or prevent pausing and arrest by RNAPII have been identified. These factors are summarized in Table 5 and a possible relationship to the RNAPII CTD is indicated.
Table 5. Transcription elongation factors

<table>
<thead>
<tr>
<th>Factor</th>
<th>Activities</th>
<th>Comments</th>
<th>Relation to RNAPII CTD</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA150</td>
<td>Promotes HIV Tat-dependent elongation</td>
<td>Regulation in a TATA box-dependent manner</td>
<td>Direct interaction with hyperphosphorylated CTD; component of RNAPII holoenzyme</td>
</tr>
<tr>
<td>CSB</td>
<td>Suppresses pausing</td>
<td>Mutated in Cockayne syndrome; required for transcription-coupled DNA repair</td>
<td></td>
</tr>
<tr>
<td>DSIF (Spt4p, Spt5p)</td>
<td>Inhibits elongation</td>
<td>Stimulates elongation under certain conditions; some mutations render yeast sensitive to 6-azauracil; genetic interactions with PPR2 and RNAPII subunit genes</td>
<td>Negative regulation of transcription through interaction with RNAPII A</td>
</tr>
<tr>
<td>ELL</td>
<td>Suppresses pausing</td>
<td>Forms complex with hSNF8, EAP20, EAP45; interacts with p53</td>
<td></td>
</tr>
<tr>
<td>ELL2</td>
<td>Suppresses pausing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elongator Elp1p, Elp2p, Elp3p</td>
<td>Histone acetyl transferase</td>
<td>Mutations render yeast sensitive to 6-azauracil and cause delays in gene activation \textit{in vivo}; genetic</td>
<td>Associated with chromatin-bound RNAPII</td>
</tr>
<tr>
<td>Elongin A</td>
<td>Suppresses pausing</td>
<td>Binds to and is stimulated by Elongin BC complex</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>--------------------</td>
<td>-------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Elongin A2</td>
<td>Suppresses pausing</td>
<td>Binds to elongin BC complex</td>
<td></td>
</tr>
<tr>
<td>FACT (Spt16p, Pob3p)</td>
<td>Promotes elongation on chromatin template</td>
<td>Involved in transcription and DNA replication; genetic interaction with SPT4 and PPR2</td>
<td></td>
</tr>
<tr>
<td>Factor 2</td>
<td>Promotes ATP-dependent transcript release</td>
<td>SWI2/SNF2 family member</td>
<td></td>
</tr>
<tr>
<td>FCP1</td>
<td>Suppresses pausing; promotes Tat-dependent elongation</td>
<td>Cooperates with TFIIF and TFII, independent of phosphatase activity; genetic interactions with PPR2, RPB1, RTF1 and TFG1</td>
<td></td>
</tr>
<tr>
<td>HMG14</td>
<td>Promotes elongation on chromatin templates</td>
<td>Might counteract histone H1 repression</td>
<td></td>
</tr>
<tr>
<td>NELF</td>
<td>Inhibits elongation</td>
<td>Possible RNA-binding activity</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Negative regulation of transcription through interaction with RNAPIIA</td>
<td></td>
</tr>
<tr>
<td>TFIIH</td>
<td>Supresses pausing</td>
<td>RAP30 and RAP74 subunits</td>
<td>RAP74 subunit stimulates CTD phosphatase</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Tat-SF1</td>
<td>Suppresses pausing and stimulates Tat-dependent elongation</td>
<td>Component of complex containing P-TEFb and hSPT5</td>
<td>Tat-SF1 complex contains RNAPIIA and can function in initiation and elongation</td>
</tr>
<tr>
<td>TFIIS</td>
<td>Prevents arrest and promotes cleavage of nascent transcript</td>
<td>Mutations render yeast sensitive to 6-azauracil; genetic interactions with genes encoding elongation factors</td>
<td></td>
</tr>
<tr>
<td>P-TEFb</td>
<td>Antagonizes NELF and DSIF and needed for Tat-dependent elongation</td>
<td>Consists of Cdk9 and cyclin T1, T2a, T2b or K</td>
<td>CTD kinase; transcription activity depends on kinase; possible yeast homolog is: CTDK-I</td>
</tr>
</tbody>
</table>
Three broad functional classes of transcription elongation factors can be defined based on their abilities to reactivate arrested RNAPII, prevent transient pausing by RNAPII or promote elongation through chromatin. One member of the first class is TFIIS. TFIIS reactivates arrested RNAPII by stimulating the catalytic center of RNAPII to cleave the nascent transcript upstream of its 3’-OH terminus. This creates a new 3’-OH terminus that is correctly positioned at the RNAPII catalytic site and can be re-extended (Wind and Reines 2000). To date, no studies have addressed the question of whether the phosphorylation state of the RNAPII CTD in any way influences TFIIS function. However, a genetic interaction between PPR2, encoding TFIIS, and FCP1 has been reported (Costa and Arndt 2000). Transcriptional arrest may also be influenced by P-TEFb, DSIF (Wada et al. 1998) and NELF.

The second class of transcription elongation factors functions by increasing the catalytic rate of transcription elongation by altering the K_m and/or Vmax of RNAPII. This class includes TFIIF (Flores et al. 1989; Price et al. 1989; Bengal et al. 1991), elongin A(SIII) and elongin A2 (Aso et al. 1995; Aso et al. 2000), the Holo-ELL complex and ELL2 (Shilatifard et al. 1997; Shilatifard 1998), CSB (Selby and Sancar 1997), CA150 (Sune et al. 1997; Sune and Garcia-Blanco 1999), Tat-Sfl (Zhou and Sharp 1996; Li and Green 1998) and possibly FCP1 (Cho et al. 1999; Lehman and Dahmus 2000).

A third class of factors regulates the rate of transcription elongation through nucleosomes and includes FACT (Facilitates Chromatin Transcription) (Orphanides et al. 1998), HMG14 (Ding et al. 1994; Ding et al. 1997) and, possibly, elongator (Wittschieben et al. 1999).
Although this is a convenient classification, overlap and crosstalk between the different classes of elongation factors exist as I will discuss below. The following discussion focuses primarily on elongation factors whose function is related to the CTD.

3.5.1 TFIIF

The interaction between TFIIF and CTD phosphatase is the main topic of Chapter 3 of this thesis, although it is not yet clear at which stage of the transcription cycle the interaction of TFIIF and Fcp1p is of functional importance. Human TFIIF is a general transcription factor, consisting of RAP30 and RAP74 subunits, which was originally purified based on its ability to bind RNAPII directly (Burton et al. 1988). TFIIF regulates initiation by inhibiting non-specific DNA-binding by RNAPII (Killeen and Greenblatt 1992) and recruiting RNAPIIA to the DNA-TFID-TFIIB complex (Flores et al. 1991). This provides a scaffold for the addition of TFIIE and TFIIH. TFIIF participates together with TFIIE in the first stage of DNA melting by inducing DNA supercoiling around RNAPII. RAP30/RAP74 can be crosslinked to promoter DNA in the pre-initiation complex both upstream and downstream of the TATA box, and this might contribute to a tightly wrapped structure of DNA around RNAPII and the GTFs (Forget et al. 1997; Fu et al. 1998; Robert et al. 1998). Subsequently, TFIIF promotes the formation of the first phosphodiester bond of the RNA transcript and also acts during very early elongation to decrease the frequency of abortive initiation (Lei et al. 1999; Ren et al. 1999; Yan et al. 1999). TFIIF remains bound to RNAPII during chain elongation, enhancing the polymerization rate and processivity (Flores et al. 1989; Price et al. 1989;
Bengal et al. 1991).

The functional regions of TFIIF and some of its interacting proteins have been determined and are diagrammed in Figure 2.
Figure 2. Domain structure of human TFIIF (RAP30/74). Interacting domains of human RAP30 and RAP74 and the biochemical function of these regions are assigned from published data. See text for details. Sites phosphorylated by the RAP74 kinase activity are also indicated (AuPS).
TFIIB binding

RAP74 binding

DNA binding, Initiation

eelongation

RNAPII binding, elongation

RAP30

RAP74

TAF250 binding

RNAPII binding

RAP30 binding, RNAPII delivery to PIC

Rearrangement of DBRNAPIIIFE complex, DNA wrapping

Accurate Initiation, elongation stimulation

Multiple round transcription stimulation
The central region of RAP30 is necessary for elongation activity and has sequence homology to bacterial σ-factor region 2, which binds to bacterial core polymerase (McCracken and Greenblatt 1991). The amino-terminal region of RAP30 comprising residues 2-98, is sufficient for RAP74-binding, while extension of RAP30 to residue 152 suffices for TFIIB binding (Fang and Burton 1996). The carboxy-terminal region of RAP30 is necessary for transcription initiation, binds nonspecifically to DNA (Tan et al. 1994) and contains a winged-helix DNA-binding motif (Groft et al. 1998). The amino-terminus of RAP74 can bind to the TAF11250 subunit of TFIID (Ruppert and Tjian 1995) and provides a minimal RAP30-binding domain. Together, interacting portions of RAP30 and RAP74 form a novel “triple barrel” dimerization fold (Gaiser et al. 2000). The amino-terminal portion of RAP74 suffices for the formation of a functional pre-initiation complex (Wang and Burton 1995; Lei et al. 1999) but increased initiation and elongation activities requires additional amino acids within the charged central part of the protein. This portion of RAP74 is important for DNA wrapping, first phosphodiester bond formation and elongation processivity (Lei et al. 1998). The charged central portion of RAP74 is likely unstructured (Yong et al. 1998; Gaiser et al. 2000). The carboxy-terminal region of RAP74 can interact with TFIIB (Fang and Burton 1996), a RNAPII/DNA complex (Wang and Burton 1995) and the CTD phosphatase FCP1 (Archambault et al. 1998b) and might mediate the role of TFIIF in multiple round transcription (Lei et al. 1998).

Human RAP74 is a phosphoprotein in vivo (Sopta et al. 1985; Kitajima et al. 1985).
1994). It can be phosphorylated \textit{in vitro} by TFIIH (Ohkuma and Roeder 1994; Yankulov and Bentley 1997), TAF\textsubscript{n}250 (Dikstein et al. 1996; Yonaha et al. 1997), casein kinase II and may also possess an autophosphorylation activity (Rossignol et al. 1999). The autophosphorylation of RAP74 at Ser385 and Thr389 by its serine/threonine kinase activity might lead to a slight reduction of the elongation activity of TFIIF (Rossignol et al. 1999). However, it has also been suggested that phosphorylation of RAP74 enhances TFIIF initiation and elongation-stimulating activities and might stabilize the TFIIF-RNAPII interaction (Kitajima et al. 1994). Phosphorylation of RAP74 by the kinase activity of the TAF\textsubscript{n}250 subunit of TFIID might also correlate with HIV Tat-activated transcription (Zhou et al. 2000b). RAP74 co-localizes with elongating RNAPII and the casein kinase II alpha subunit on polytene chromosomes of \textit{Chironomus} salivary glands in a DRB-sensitive manner (Egyhazi et al. 1999).

Interestingly, yeast TFIIF contains an additional subunit besides the homologs of human RAP30 and RAP74 (Henry et al. 1994). This additional TFIIF subunit is also found in TFIID (Henry et al. 1994) and the SWI/SNF complex involved in chromatin remodeling (Cairns et al. 1996a). Very little is known about the activities of yeast TFIIF.

\subsection*{3.5.2 CTD phosphatase as an elongation factor}

A role for FCP1 in the regulation of early and processive transcription elongation complexes is slowly emerging from \textit{in vitro} and \textit{in vivo} studies. Both forms of human FCP1, the active CTD phosphatase FCP1 (Cho et al. 1999) and amino-terminally truncated FCP1a (Archambault et al. 1998b), which does not have CTD phosphatase
activity, can synergize with TFIIF and TFIIS to enhance transcription elongation in vitro, suggesting that this function is independent of CTD phosphatase activity (Cho et al. 1999). In yeast, strains with a carboxy-terminally truncated version of Fcp1p encoded by the fcp1-110 allele are synthetically lethal with a deletion of the rfl gene that likely encodes a transcription elongation factor (Costa and Arndt 2000). A strain with this fcp1-110 allele and a deletion of the ppr2 gene, which encodes TFIIS, exhibit a strong growth defect and enhanced inositol sensitivity compared to the fcp1-110 strain and the ppr2Δ strain (Costa and Arndt 2000). In addition, the fcp1-110 mutation causes strains to be weakly sensitive to high concentrations of 6-azauracil, an inhibitor of transcription elongation (Costa and Arndt 2000). Therefore, both yeast and human FCP1 are likely to be RNAPII elongation factors.

Recombinant human FCP1 is active as a CTD phosphatase in ternary elongation complexes obtained by Sarkosyl treatment of in vitro transcription reactions assembled from purified factors (Cho et al. 1999). In contrast to these studies, experiments performed in vitro using a transcription system derived from HeLa nuclear extracts suggest that CTD phosphatase activity is influenced by factors present in nuclear extracts and depends on the length of the RNA transcript (Lehman and Dahmus 2000). Elongation complexes 11-24 nucleotides downstream of the transcription start site are more sensitive to CTD dephosphorylation than complexes 50-150 nucleotides downstream (Lehman and Dahmus 2000). Therefore, a conformational change might occur in RNAPII as it clears the promoter, resulting in a reduced ability of CTD phosphatase to dephosphorylate the
CTD. This is consistent with earlier data demonstrating that RNAPII goes through several conformational transitions during early elongation (Samkurashvili and Luse 1998).

The HIV-1 transactivator protein Tat, which stimulates elongation by RNAPII, can interact with FCP1 (J. Archambault and J. Greenblatt, unpublished data) and regulate its CTD phosphatase activity (Marshall et al. 1998). HIV Tat inhibits CTD phosphatase activity on free RNAPII and also on RNAPII in early elongation complexes initiated from the promoter in the HIV-1 LTR (Marshall et al. 1998; Marshall and Dahmus 2000). Paused RNAPII elongation complexes initiated from this promoter become more resistant to CTD phosphatase activity after passing nucleotide +25 (Marshall and Dahmus 2000). One limitation of these studies is the fact that highly purified CTD phosphatase from HeLa cells was used rather than recombinant FCP1. Therefore, FCP1-associating factors present in the preparation might influence the experimental outcome (Marshall et al. 1998; Lehman and Dahmus 2000; Marshall and Dahmus 2000).

3.5.3 DSIF (Spt4p/Spt5p), NELF and P-TEFb

Two classes of elongation factors responsible for DRB-sensitive transcription have been identified. The class of negatively acting factors includes DSIF (DRB-sensitivity inducing factor) (Wada et al. 1998) and NELF (negative elongation factor) (Yamaguchi et al. 1999a). P-TEFb is a positively acting, DRB-sensitive CTD kinase (described in more detail above). These three factors act together to regulate transcription elongation.

Human DSIF is a heterodimer consisting of 160kDa and 14kDa subunits (Wada et
al. 1998). An additional co-factor, NELF, cooperates with DSIF to strongly inhibit RNAPII elongation (Yamaguchi et al. 1999a). NELF is composed of five polypeptides, the smallest of which is identical to RD, a putative RNA-binding protein named after its Arg-Asp repeat (Yamaguchi et al. 1999a). The other four subunits of NELF have not yet been described (Yamaguchi et al. 1999a). DSIF and NELF bind concomitantly to hypophosphorylated RNAPIIA and negatively regulate elongation complexes with nascent transcripts more than 40 nucleotides long (Yamaguchi et al. 1999a). DSIF/NELF can even inhibit RNAPII activity in a non-specific elongation assay *in vitro* in the absence of any other factors (Yamaguchi et al. 1999a). In this assay, both RNAPIIA and RNAPIIB that lacks the CTD are inhibited, suggesting that the CTD is not necessary for DSIF/NELF function. However, the elongation activity of RNAPIIO carrying a hyperphosphorylated CTD is not inhibited. Therefore, DSIF/NELF interacts with a region of RNAPII outside of the CTD, but phosphorylation of the CTD might block the interaction of DSIF/NELF with RNAPIIO (Yamaguchi et al. 1999b).

Importantly, P-TEFb can overcome the negative effect of DSIF/NELF by phosphorylating RNAPII in early elongation complexes and, by doing so, promotes transcriptional elongation (Yamaguchi et al. 1999a). Therefore, a model has been proposed to explain the action of DSIF/NELF and P-TEFb. Processive transcription elongation that is insensitive to DRB takes place in a highly purified transcription system containing RNAPII and the general transcription factors, because both negative and positive elongation factors are not present. In a more crude system, DSIF/NELF bind to RNAPIIA and inhibits elongation. However, P-TEFb can phosphorylate the CTD,
promoting the conversion from RNAPIIA to RNAPII. This prevents the functioning of DSIF/NELF, possibly by inducing a conformational change that masks the docking site for DSIF/NELF on RNAPII. Dissociation of DSIF/NELF from RNAPII then allows productive elongation to occur. When DRB is present, the P-TEFb kinase is inhibited, allowing DSIF/NELF to continue inhibiting RNAPII elongation.

The elongation factor FACT (Spt16p/Pob3p), which I describe in more detail below, functions in conjunction with P-TEFb to alleviate transcription inhibition by DSIF/NELF in a minimal system (Wada et al. 2000). The CTD kinase activity of TFIIH is dispensable for this cooperative effect of FACT and P-TEFb, demonstrating that TFIIH-dependent CTD phosphorylation is not involved in the regulation of FACT and DSIF/NELF activities (Wada et al. 2000). It is not clear whether the additional effect of FACT on DSIF/NELF is in any way connected to the RNAPII CTD. P-TEFb can also phosphorylate residues within the carboxy-terminal domain of the hSPT5 subunit of DSIF, and this is important for hSPT5 function (Ivanov et al. 2000), suggesting yet another level of regulation.

Studies in yeast and Drosophila support a role for DSIF in transcription elongation in vivo. The yeast homologs of the human DSIF subunits, Spt4p and Spt5p, also form a complex. SPT4 and SPT5, along with SPT6, were initially identified in a screen for suppressors of cis- and trans-acting mutations that affect promoter function (Winston and Carlson 1992). The three proteins Spt4p, Spt5p and Spt6p appear to be involved in a common function in vivo (Hartzog et al. 1998). Some mutations in spt4 and spt5 cause yeast to be sensitive to the drug 6-azauracil that inhibits transcriptional elongation by
reducing the intracellular concentration of GTP (Hartzog et al. 1998). Several connections exist between SPT5 and the genes encoding the two largest subunits of RNAPII, RPB1 and RPB2, and Spt5p can associate with RNAPII in vivo (Hartzog et al. 1998). This is interesting because Spt5p contains a region of sequence similarity with the transcription elongation factor NusG from Escherichia coli that interacts with E.coli RNAP (Li et al. 1992; Hartzog et al. 1998). Additionally, a yeast strain with a mutated version of spt5 and a deletion of the ppr2 gene which encodes the transcription elongation factor TFIIS has a much more severe growth defect than strains with only one of these mutations (Hartzog et al. 1998). Most convincingly, the Drosophila homologs of Spt5p and Spt6p co-localize extensively with hyperphosphorylated, actively elongating RNAPIIO to transcriptionally active sites during salivary gland development and to heat shock loci upon heat shock, but do not co-localize with non-transcribing RNAPIIA (Andrulis et al. 2000; Kaplan et al. 2000). Spt4p, Spt5p and Spt6p might also have a role in recombination and chromosome segregation (Basrai et al. 1996; Malagon and Aguilera 1996), suggesting the possibility that these proteins establish certain chromatin states necessary for a variety of cellular functions.

3.5.4 hSPT5, Tat-Sf1, RAP30 and Tat-SF

Human SPT5 together with TAT-Sf1 is required for Tat activation in nuclear extracts (Wu-Baer et al. 1998; Kim et al. 1999). A small fraction of hSPT5 and RNAPII is associated with TAT-Sf1 and the RAP30 subunit of TFIIF but not RAP74 (Kim et al. 1999). Hence, TAT-Sf1 and hSPT5 might interact with RAP30 to control elongation and
there is evidence that TAT-Sfl is a general elongation factor (Li and Green 1998). However, the TAT-SF complex that supports activated transcription by HIV Tat contains Tat-Sfl, human SPT5, P-TEFb and RNAPIIA (Parada and Roeder 1999), suggesting that TAT-SF may function in transcription initiation although it does not have any Srb/mediator proteins.

3.5.5 FACT (Spt16p/Pob3p)

FACT was identified and purified as a factor that overcame a nucleosomal block to RNAPII in vitro. It is a heterodimeric factor comprised of a human homolog of yeast Cdc68p/Spt16p (FACTp140) and the HMG1-like protein SSRP1 (FACTp80), the human homolog of yeast Pob3p. Proteins containing HMG-1 domains bind to DNA at the cross-over points where it enters and exits a nucleosome (Bustin and Reeves 1996). FACT specifically interacts with nucleosomes and histone H2A/H2B dimers and might therefore promote nucleosome disruption during transcription elongation. Nucleosomes containing transcribed sequences are deficient in histones H2A and H2B and are preferentially bound by RNAPII (Baer and Rhodes 1983). FACT activity in vitro is abrogated by covalent crosslinking of nucleosomes that prevents the disruption of the histone octamer indicating that FACT might bind to nucleosomes and displace H2A/H2B dimers to facilitate transcription elongation (Orphanides et al. 1998).

Yeast SPT16 is genetically linked to the genes encoding the RNAPII transcription elongation factors TFIIS and Spt4p (Orphanides et al. 1999). Mutations in spt16 partly reverse the extreme sensitivity of ppr2Δ and spt4 strains to the transcription elongation
inhibitor 6-azauracil (Orphanides et al. 1999). The Spt16p/Pob3p complex is an abundant nuclear complex in yeast that was originally isolated based on an interaction with DNA polymerase α. Consistent with this, the Spt16p/Pob3p complex has a role not only in transcription but also in DNA replication (Schlesinger and Formosa 2000). Independent of its role in transcription elongation on chromatin templates, but consistent with the genetic interaction between SPT16 and SPT4 (Orphanides et al. 1999), FACT cooperates with P-TEFb to relieve transcriptional inhibition by DSIF (Spt4p/Spt5p) and NELF on naked transcription templates (Wada et al. 2000).

3.5.6 CA150

CA150 binds a CTD preparation that is phosphorylated by CTDK-I, which presumably mimics the phosphorylation that occurs in elongating RNAPII (Carty et al. 2000). The main binding site in CA150 for the CTD is located within its FF domain, named for the presence of two conserved Phe residues, rather than in its WW domain (Carty et al. 2000). Interestingly, CA150 was previously identified in a fraction required for transcriptional activation by HIV Tat and can associate with a human RNAPII holoenzyme complex (Sune et al. 1997). It is not clear how this can be reconciled with the selective binding of CA150 to phosphorylated CTD that has been reported. Human CA150 has a low level of homology to the Gal11p and Sin4p components of the yeast mediator (Sune et al. 1997). A role for CA150 in transcriptional elongation that was dependent on a functional TATA box has been suggested (Sune and Garcia-Blanco 1999),
but it should be noted that an elongation-stimulating activity of CA150 has not been
directly demonstrated. The effects of CA150 on HIV Tat activated transcription could
reflect a role of CA150 in mRNA processing.

3.5.7 An elongating form of RNAPII

A form of RNAPII involved in transcriptional elongation was purified from yeast
chromatin. This elongating RNAPII complex was distinct from the RNAPII holoenzymes
previously described because it lacked Srb/mediator proteins and instead contained a
novel multi-subunit complex termed elongator (Otero et al. 1999). The stable association
of elongator with RNAPII requires a hyperphosphorylated CTD (Otero et al. 1999).
Therefore, a mediator-elongator cycle during transcription that depends on the
phosphorylation state of the CTD has been proposed (Svejstrup et al. 1997). Mediator is
associated with RNAPIIA during transcription initiation and elongator with RNAPIIO
during transcriptional elongation.

The ELP1, ELP2 and ELP3 genes encoding the elongator subunits are not essential
(Otero et al. 1999; Wittschieben et al. 1999; Fellows et al. 2000). Strains lacking these
genes are slow in adapting to new growth conditions and sensitive to high salt
concentrations (Otero et al. 1999; Wittschieben et al. 1999; Fellows et al. 2000). These
phenotypes might be a consequence of defects in the transcription of genes required for
cells to adapt to changing conditions. Consistent with a role for elongator in
transcriptional elongation in vivo, strains with deletions of the elp genes and the ppr2 gene
encoding the transcription elongation factor TFIIS are more sensitive to the drug 6-
azauracil than strains with the *elp* or *ppr2* deletions alone (Otero et al. 1999; Wittschieben et al. 1999).

Elongator contains a histone acetyltransferase activity (HAT) capable of acetylating all four histones *in vitro* (Wittschieben et al. 1999). Elongator HAT activity is intrinsic to Elp3p, which is a highly conserved protein in eukaryotes (Wittschieben et al. 1999). Mutations in the tails of histone H3 and H4 cause *elp3* strains to become extremely sick, suggesting a role for elongator in chromatin remodeling *in vivo* (Wittschieben et al. 2000). Genetic studies suggest that elongator and the SAGA complex that contains the Gcn5p HAT activity might have overlapping functions *in vivo* (Wittschieben et al. 2000). The synthetic phenotypes of a *elp3Δ gcn5Δ* strain can be suppressed by concomitant mutations in the *hda1* and *hos2* genes that encode histone deacetylases (Wittschieben et al. 2000).

### 3.5.8 Linking transcription elongation, DNA repair and recombination

Transcription elongation might be linked to several DNA metabolic processes, including transcription-coupled DNA repair, homologous recombination, and maintenance of genome stability. Several proteins that have been connected to RNAPII also have roles in either DNA repair or recombination.

The CSB protein functions *in vitro* as an elongation factor by stimulating the overall elongation rate (Selby and Sancar 1997) and interacts directly with transcribing RNAPII (Tantin et al. 1997). CSB is mutated in certain patients with Cockayne
syndrome and is required in transcription-coupled nucleotide excision repair of damaged DNA (Conaway and Conaway 1999). Mutations in the hpr1 and tho2 genes of S. cerevisiae induce a hyper-recombination phenotype (Aguilera and Klein 1990; Piruat and Aguilera 1998) that can be suppressed by insertion of transcription termination signals upstream of recombinogenic regions of transcribed DNA (Prado et al. 1997).

Additionally, transcription elongation by RNAPII seems to be impaired by hpr1 and tho2 deletions (Piruat and Aguilera 1996; Prado et al. 1997) and, conversely, the hyper-recombination phenotype of these strains is enhanced when cells are grown in 6-azauracil (Chavez and Aguilera 1997). Hpr1p and Tho2p form a complex with Mtf1p and a novel protein called Thp2p that might connect transcriptional elongation with the incidence of mitotic recombination (Chavez et al. 2000). Based on these observations, it has been suggested that inappropriately paused elongation complexes in hpr1 and tho2 mutant strains provide a signal for the recruitment of the recombination machinery, leading to the hyper-recombination phenotype (Piruat and Aguilera 1998).

However, the hyper-recombination phenotype of hpr1 and tho2 strains can also be suppressed by mutations in genes encoding RNAPII subunits, the Srb/mediator component Srb2p, the general transcription factor TFIIB and the positive transcription factor Soh1p (Fan et al. 1996; Piruat and Aguilera 1996), a component of the SMCC and PC2 co-activator complexes (Malik and Roeder 2000). This could suggest that events mediated at the promoter by the Srb/mediator complex may affect the transcribing RNAPII complex. In addition, Hpr1p is also a component of the Paf1p-RNAPII holoenzyme described before (Chang et al. 1999).
3.6 Connecting CTD phosphorylation to mRNA maturation

Transcription generates the initial gene transcript, which then is converted into a functional mRNA during RNA processing. Recent genetic and biochemical studies have supported findings from early experiments connecting transcription with capping, splicing and polyadenylation. Interestingly, the CTD of RNAPII seems to be involved in the coordination of all three processing steps with transcription (Hirose and Manley 2000; Proudfoot 2000).

A number of early experiments indicated that any one of these reactions could enhance some aspect of another. The 5’cap is involved in allowing efficient transcription, splicing and polyadenylation (Jove and Manley 1982; Edery and Sonenberg 1985; Hart et al. 1985). Splicing signals present on nascent pre-mRNAs can also influence polyadenylation (Niwa et al. 1990; Niwa and Berget 1991). Mutation of the poly(A) signal of a gene not only prevents polyadenylation but also causes RNAPII to continue to transcribe well beyond the normal termination region (Whitelaw and Proudfoot 1986; Logan et al. 1987; Connelly and Manley 1988). Thus, RNA-processing events can directly influence transcriptional elongation. Transcription termination often occurs several kilobases downstream from the site of polyadenylation, and thus the molecular explanation for these early observations remained elusive until recently.

Major insight into the connection between the CTD and RNA processing has come from protein-protein interaction studies (summarized in Table 6). Table 6 also lists CTD-interacting proteins that might have functions in mRNA synthesis unrelated to processing (e.g. CA150, Rsp5p) and are discussed in other sections of Chapter 1.
Proteins are listed only if they interact directly with the CTD in protein-protein interaction assays.
<table>
<thead>
<tr>
<th>Protein</th>
<th>Binding to RNAPII</th>
<th>Binding to RNAPII0</th>
<th>CTD Binding Domain</th>
<th>Functional Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPSF (human)</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>Polyadenylation factor</td>
</tr>
<tr>
<td>CstF-50 (human)</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>Subunit of heterotrimeric polyadenylation factor</td>
</tr>
<tr>
<td>Ceg1p (yeast)</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>5' capping; guanylyltransferase</td>
</tr>
<tr>
<td>Abd1p (yeast)</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>5' capping; guanine N7 methyltransferase</td>
</tr>
<tr>
<td>Pta1p (yeast)</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>3' processing; polyadenylation factor</td>
</tr>
<tr>
<td>Prp40p (yeast)</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>Splicing factor</td>
</tr>
<tr>
<td>SRcyp/SCAF10 (human)</td>
<td>N.D.</td>
<td>+</td>
<td>RS</td>
<td>Peptidylprolyl-cis-trans-isomerase and RS domain</td>
</tr>
<tr>
<td>Nrd1p (yeast)</td>
<td>+</td>
<td>+</td>
<td>CID</td>
<td>Related to human SCAF8 and SCAF4; contains RS and RRM; interacts with Nab1p; function in RNA processing</td>
</tr>
<tr>
<td>SCAF1 (human)</td>
<td>-</td>
<td>+</td>
<td>CID</td>
<td>Similarity to SR protein family; SCAF8 also contains RRM</td>
</tr>
<tr>
<td>SCAF8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>Ess1p (yeast)</td>
<td>-</td>
<td>+</td>
<td>WW</td>
<td>Peptidylprolyl-	extit{cis-trans-} isomerase; involved in transcription termination (yeast)</td>
</tr>
<tr>
<td>PIN1 (human)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA150 (human)</td>
<td>-</td>
<td>+</td>
<td>FF</td>
<td>Role in transcription elongation; component of RNAPII holoenzyme</td>
</tr>
<tr>
<td>Rsp5p (yeast)</td>
<td>+</td>
<td>Reduced</td>
<td>WW</td>
<td>E3 ubiquitin ligase; protein degradation</td>
</tr>
</tbody>
</table>

N.D. = not determined

CID = CTD interacting domain, not overlapping with RS domain

RRM = RNA recognition motif

RS = Arg-Ser diamino acid motif
3.6.1 Capping

The cap structure found at the 5’ end of all eukaryotic mRNAs is formed before the transcript has reached a size of ~30 nucleotides (Coppola and Luse 1984; Jove et al. 1984). Capping is carried out by a series of three enzymatic activities (Shuman 1995; Shuman 2000). First, the γ-phosphate of the first nucleotide of the pre-mRNA is removed by a RNA triphosphatase. This is followed by transfer of GMP to the resulting diphosphate end by RNA guanylyltransferase. In the last step, a methyl group is added to the N7 position of the cap guanine by RNA (guanine-7-) methyltransferase to form the m7G(5’ppp(5’)N cap. Capping enzyme consists of a heterodimer of RNA triphosphatase (Cet1p) and RNA guanylyltransferase (Ceg1p) in S. cerevisiae. but these activities are part of a bifunctional protein in metazoans. RNAPII can bind directly and selectively to mammalian capping enzyme, and this seems to involve phosphorylated CTD and the guanylyltransferase domain of the capping enzyme (McCracken et al. 1997a; Yue et al. 1997; Ho et al. 1998). The yeast capping enzymes Ceg1p and the RNA (guanine-7-) methyltransferase (Abd1p) can bind directly in vitro to phosphorylated CTD but not to non-phosphorylated CTD (McCracken et al. 1998; Rodriguez et al. 2000) and conditional mutations in the genes encoding the capping enzymes are synthetically lethal with truncation alleles of the CTD in vivo (Rodriguez et al. 2000). RNAs synthesized by RNAPII carrying a shortened CTD undergo inefficient capping in transfected mammalian cells (McCracken et al. 1997a).

The CTD not only binds to capping enzyme but also seems to regulate allosteric
activation of the guanylyltransferase activity in a manner dependent on a differentially phosphorylated CTD. The first step of guanylyltransferase reaction involves the formation of a covalent enzyme-GMP intermediate. This intermediate cannot be detected in association with a phosphorylated GST-CTD - Ceglp complex (Cho et al. 1997). In fact, the phosphorylated CTD inhibits Ceglp activity, but this can be reversed, and actually enhanced, by addition of the Cetlp triphosphatase (Cho et al. 1998). Therefore, Ceglp activity might be allosterically regulated by interactions with RNAPIIO and Cetlp, possibly coordinating guanylylation with triphosphatase activity (Cho et al. 1998). Although allosteric regulation of the mammalian capping enzyme also occurs through interaction with the phosphorylated CTD, the situation is slightly different.

Synthetic CTD peptides containing phospho-Ser at either Ser2 or Ser5 of the YSP[TSP[S heptapeptide repeat can bind the isolated guanylyltransferase domain of mouse capping enzyme (Ho and Shuman 1999). This binding is dependent on phosphorylation (Ho and Shuman 1999). Whereas CTD peptides containing phospho-Ser5 stimulate formation of the enzyme-GMP intermediate, CTD peptides containing phospho-Ser2 have no effect on either basic enzyme activity or on guanylylation activated by phospho-Ser5 (Ho and Shuman 1999). Thus, mammalian capping enzyme seems to contain two independent binding sites for the phosphorylated CTD: one acts as an allosteric activator site and binds to Ser5 of the CTD; the other is specific for Ser2 and does not activate. Studies in yeast also support the importance of phosphorylation of Ser5. A set of conditional mutations in rpb1 in which Ser5, but not Ser2, of the CTD repeats is replaced with Ala is synthetically lethal with a cegl mutation (Rodriguez et al. 2000).
The CTD kinase activity of TFIIH is probably responsible for recruiting capping enzyme to RNAPII. GST-CTD preparations phosphorylated by yeast Kin28p-Ccl1p, Srb10p-Srb11p or CTDK1 can bind capping enzyme in vitro (Rodriguez et al. 2000). However, strains with mutations in cegl only have synthetic growth defects in kin28 mutant backgrounds and not in strains with ctkl or srb10 mutations (Rodriguez et al. 2000). Most importantly, Kin28p phosphorylates the CTD specifically on Ser5 in vivo during promoter clearance and regulates the association of the capping enzyme components with active genes (Komarnitsky et al. 2000). Chromatin immunoprecipitation experiments have shown that Cegl1p and Abd1p cross-link in vivo to the 5' ends of genes, and this localization requires the CTD (Schroeder et al. 2000). Cegl1p is released early in elongation but Abd1p can travel with elongating RNAPII as far as the 3' end of a gene (Komarnitsky et al. 2000; Schroeder et al. 2000). Kin28p-mediated Ser5 phosphorylation also decreases as the transcription complex moves from the 5' towards the 3'end of a gene, coinciding with the release of Cegl1p (Komarnitsky et al. 2000; Schroeder et al. 2000). Whereas Kin28p is required for the association of capping enzymes with transcription complexes, the CTD phosphatase Fcp1p, which is the subject of this thesis, may be necessary for the dissociation of capping enzymes from the elongation complex (Schroeder et al. 2000).

Another interesting connection between mRNA capping and transcription elongation is the physical and functional interaction of human SPT5, a component of DSIF, with the human capping enzyme (Wen and Shatkin 1999). Human SPT5 binds human capping enzyme and stimulates guanylylation and mRNA capping by severalfold
(Wen and Shatkin 1999). However, simultaneous addition of phosphorylated GST-CTD protein and SPT5 has no additive effect, possibly pointing to a redundant function of the two proteins (Wen and Shatkin 1999). Yeast Spt5p has not yet been shown to be involved in the capping reaction.

In summary, it seems possible that capping marks the complete switch from transcription initiation to elongation. Because RNAPII stalls on certain genes soon after transcriptional initiation, it has been suggested that capping could play a role in transcription attenuation (Proudfoot 2000). Capping could be used to regulate gene expression by releasing the polymerase from its elongation block.

3.6.2 Splicing and transcription

Splicing of introns is the next RNA processing reaction to occur on the nascent transcript. It takes place in a large complex, called the spliceosome, that is composed of small nuclear ribonucleoprotein particles (snRNPs) and non-snRNP proteins, including members of the Ser/Arg-rich (SR) protein family (Kramer 1996; Manley and Tacke 1996). The complexity of splicing increases dramatically from yeast to mammals. In budding yeast, only 3% of all genes contain intervening sequences, and those are usually small and located near the 5’ end of the transcript. However, the intervening sequences are overrepresented in highly expressed genes, such as those encoding ribosomal proteins, and thus 30% of the total primary transcript pool contains introns (Ares et al. 1999; Lopez and Seraphin 1999). In mammals, however, most genes are highly intronic, with well over 90% of the genome consisting of non-coding regions.
Almost all evidence linking transcription and splicing comes from metazoans. Cytological studies on chromosomal spreads from *Drosophila* showed first that splicing can occur co-transcriptionally (Beyer and Osheim 1988; Bauren et al. 1998). Factors required for splicing can be found localized at sites of active transcription in mammalian nuclei (Zhang et al. 1994). Consistent with this, RNAPII, but not RNAPII A, is present in active spliceosomes (Chabot et al. 1995; Mortillaro et al. 1996; Yuryev et al. 1996; Kim et al. 1997). Splicing reactions *in vitro* can be inhibited by antibodies directed against the CTD (Chabot et al. 1995; Yuryev et al. 1996). Conversely, CTD peptides that are overexpressed *in vivo* become phosphorylated and inhibit splicing (Du and Warren 1997).

In cultured mammalian cells, RNAPII with a truncated CTD produces transcripts that are inefficiently spliced (McCracken et al. 1997b). This has been confirmed by microscopic observations showing co-localization of a reporter transcript and splicing factors when the transcript was produced by RNAPII with a full length CTD, but no co-localization when the transcript was produced by RNAPII with a truncated CTD (Misteli and Spector 1999).

Purified RNAPII can strongly activate the splicing of several different pre-mRNAs in a reconstituted splicing system, suggesting that the RNAPII plays a direct and active role in splicing *in vitro* in the absence of transcription (Hirose et al. 1999). RNAPII might stimulate splicing by accelerating the rate of one of the first steps in spliceosome assembly. In contrast, RNAPII A can inhibit splicing through inhibition of the formation of early pre-splicing complexes (Hirose et al. 1999). Experiments with CTD-less RNAPII B and a GST-CTD fusion protein show that the CTD is necessary but
not sufficient for the effects of RNAPIIO on splicing [Hirose, 1999 #3682. Importantly, the WW domain-containing splicing factor Prp40p from yeast can interact with the phosphorylated CTD, providing the first example of a direct splicing factor – RNAPII interaction (Morris and Greenleaf 2000). Prp40p is involved in bridging the 5' and 3' splice sites during the commitment stage of spliceosome assembly, thereby implicating the phosphorylated CTD in the commitment stage.

The interaction between the CTD and the spliceosome might also be mediated by other splicing factors, including SR proteins that interact with the CTD through a specific CTD-interacting domain (CID). These are the SR-like CTD-associated factors (SCAFs) or CTD-associated SR-like proteins (CASP) (Corden and Patturajan 1997). The first four SCAFs were originally identified in a two-hybrid screen for CTD-interacting proteins (Yuryev et al. 1996). SCAFs generally contain one or two RNA-binding domains (RNA recognition motifs : RRM), an Arg-Ser-rich region (the SR domain) and a CTD-interacting domain (CID) (Corden and Patturajan 1997). Human SRcyp/SCAF10 has a cyclophilin-like peptidylprolyl-cis-trans-isomerase (PPIase) domain at its amino-terminus and an Arg-Ser-rich region at its carboxy-terminus that is involved in binding the CTD (Bourquin et al. 1997). SRcyp co-localizes with the splicing factor SC35, itself a SR protein, and with the snRNP protein U1-70K (Bourquin et al. 1997). SCAF8 interacts only with phosphorylated CTD or RNAPII and can localize in cells to foci that overlap with sites for transcription and processing (Patturajan et al. 1998b). This provides indirect evidence for a role of the SCAF proteins in coupling transcription and splicing. Nrd1p, a yeast protein with similarity to SCAF proteins, including a CID, an SR-like
region that is very uncommon for yeast, and an RBD, may have some function in mRNA processing (Steinmetz and Brow 1996; Steinmetz and Brow 1998). Nrdlp is thought to regulate mRNA abundance by binding to a gene-specific cis-acting element (Steinmetz and Brow 1996) and has been genetically and physically linked to the CTD, the hnRNP protein Nab3p and the CTD kinase CTDK-I (Conrad et al. 2000). Although this evidence supports a role for Nrdlp on the interface of transcription and mRNA processing, no evidence points to a direct involvement in splicing.

Surprising additional support for a role of SR proteins in the coupling of transcription and splicing has come from studies showing that different promoters can dictate the splicing patterns of transcripts initiated on those promoters. An alternatively spliced intron from the fibronectin gene was used to demonstrate that one intron was included or excluded depending on the promoter used for transcription (Cramer et al. 1997). Importantly, overexpression of various SR proteins also affected the splicing pattern and sometimes even antagonized the promoter effects (Cramer et al. 1999). It was suggested that these results reflect promoter-dependent association of SR proteins with the CTD (Proudfoot 2000), perhaps controlled by differential phosphorylation.

Besides the SR proteins, other proteins might be involved in the functional coupling of transcription and splicing (Hirose and Manley 2000). These include TLS/FUS-related proteins, such as the TFIID component TAF\textsubscript{n}68, or the Ewing’s sarcoma protein (EWS) which interacts through its amino-terminal domain with a hyperphosphorylated RNAPII complex and recruits SR splicing factors (Yang et al. 2000).
3.6.3 Polyadenylation and transcription termination

Polyadenylation of the mRNA transcript is a two-step process in which endonucleolytic cleavage of the pre-mRNA is followed by poly(A) addition to the 3' end of the upstream cleavage product. Multiple protein factors are involved in this process (Colgan and Manley 1997; Zhao et al. 1999). These include cleavage/polyadenylation specificity factor (CPSF), cleavage stimulation factor (CstF), the two cleavage factors CFI and CFII, and poly(A) polymerase. CPSF in mammalian cells is associated with the general transcription factor TFIID and seems to be transferred to RNAPII at the onset of transcriptional elongation (Dantonel et al. 1997). CPSF and CstF are also found in certain RNAPII holoenzyme preparations, and both interact directly with a GST-CTD fusion protein in a phosphorylation-independent manner (McCracken et al. 1997b). As is the case with capping and splicing, RNAs synthesized by RNAPII carrying a shortened CTD are inefficiently polyadenylated in transfected culture cells (McCracken et al. 1997b).

RNAPII is directly involved in the polyadenylation reaction in the absence of transcription. Purified RNAPIIA and RNAPIIO can activate 3’ cleavage, which is the first step of polyadenylation, in a reconstituted system (Hirose and Manley 1998). Cleavage is also efficiently stimulated by unphosphorylated GST-CTD and hyperphosphorylated GST-CTD fusion proteins with the latter form being more active (Hirose and Manley 1998). Hence, in contrast to the splicing reaction described above, the CTD is both necessary and sufficient for activation of polyadenylation. Immunodepletion of RNAPII from nuclear extracts caused an inhibition of 3’ cleavage that could be rescued by adding back purified RNAPII (Hirose and Manley 1998).
Several lines of evidence imply that there is a close connection between transcription and 3’end processing in yeast. Transcription of the *CUP1* gene by RNAPII with a shortened CTD leads to defects in 3’ end processing, whereas the *CYC1* and *YHR54* transcripts are efficiently polyadenylated (McNeil et al. 1998). This suggests that the coupling between transcription and polyadenylation might be gene-specific in yeast. Evidence for a physical connection is provided by the binding of the polyadenylation factor Pta1p to phosphorylated GST-CTD (Rodriguez et al. 2000). Another polyadenylation factor, Hrp1p, which is a component of the cleavage factor IB in yeast, can be crosslinked *in vivo* to promoter regions and also throughout the coding regions of genes, suggesting that it associates with initiating and elongating RNAPII (Komarnitsky et al. 2000). Most importantly, the peptidylprolyl-cis-trans-isomerase Ess1p binds to the hyperphosphorylated CTD (Morris et al. 1999) and mutations in *ess1* were uncovered in a genetic screen for defects in 3’end formation (Hani et al. 1999). I will discuss Ess1p in more detail below.

Termination of transcription in yeast following the poly(A) signal of the *CYC1* gene depends on the signal and on 3’ cleavage factors but not polyadenylation factors. Yeast strains with temperature-sensitive mutations in genes encoding 3’ cleavage factors fail to terminate transcription at the non-permissive temperature. Thus, RNA cleavage might trigger RNAPII termination, possibly due to the highly unstable nature of the remaining downstream RNA that lacks a 5’ cap or triphosphate. According to this model, the action of a hypothetical 5’ to 3’ exonuclease would then degrade the RNA and cause RNAPII to terminate (Proudfoot 2000). Another model for termination suggests that a
modification occurs in RNAPII as it passes the polyadenylation signal, and this causes the enzyme to be less processive and more likely to terminate (Proudfoot 2000). This model is supported by several studies showing that cleavage of the transcript is not apparent prior to the site of termination. For example, termination of transcription initiated on minichromosomes in *Xenopus* oocytes depends on a functional poly(A) signal, but the nascent transcript is not cleaved before termination (Osheim et al. 1999).

However, these two models might actually converge. In transfected mammalian cells, termination occurs ~1000bp downstream of the polyadenylation site, and mRNAs that extend to the apparent termination site are not cleaved (Dye and Proudfoot 1999). Similar results were obtained by analyzing nascent transcripts from the heavily transcribed Balbiani ring 1 gene isolated from salivary glands of the dipteran *Chironomous tentans*. This means that RNAPII elongates into the 3’ flanking region of the gene up to the site of termination before 3’ cleavage occurs at the poly(A) site (Bauren et al. 1998).

A G-rich pause site can enhance polyadenylation at an upstream poly(A) site in a coupled processing/transcription reaction (Yonaha and Proudfoot 1999) and this supports the requirement for RNAPII, and specifically the CTD, in 3’ end processing. One possibility is that RNAPII stalled at a pause site stimulates 3’ cleavage and that 3’ cleavage is followed by endonucleolytic degradation and transcription termination. It is further possible that recognition or processing of the terminal intron is involved in the activation of polyadenylation and ultimately in transcription termination. (Dye and Proudfoot 1999)

These studies show that the RNAPII CTD has an important role in 3’end
processing and transcription termination, but it is not clear whether CTD
dephosphorylation by the CTD phosphatase FCP1 described in this thesis contributes to
these processes.

3.6.4 Compartmentalization of transcription complexes

Localization of transcription and RNA-processing proteins in cells has provided
visual evidence for the connection of these processes, supporting the biochemical and
genetic evidence that I described above. Mammalian cells contain 20-40 nuclear speckles
that appear to correspond to interchromatin granule clusters (Misteli and Spector 1998).
Speckles seem to be storage or assembly sites for splicing complexes and contain
hyperphosphorylated RNAPII and polyadenylation factors at their periphery
(Bregman et al. 1995; Schul et al. 1998). During transcriptional activation, RNAPII and
the SR proteins ASF/SF1 relocate from speckles to the sites of active transcription
(Bregman et al. 1995; Misteli et al. 1997). The relocation of splicing factors depends on
the CTD of RNAPII, consistent with the biochemical studies (Misteli and Spector 1999).

In *Xenopus* oocytes, all three nuclear RNA polymerases and associated factors
such as TFIIF, TFIIIS, splicing factors, and U7snRNP are initially accumulated in nuclear
structures called Cajal bodies (Gall et al. 1999). RNAPII and transcription and processing
factors are subsequently transported to the sites of RNA synthesis as interchromatin
granule clusters containing multiple, large transcriptosomes with all the factors necessary
for synthesis of the mature mRNA (Gall et al. 1999). Interestingly, these structures
contain mostly hyperphosphorylated RNAPII and the question of dephosphorylation
has not been addressed. Splicing and polyadenylation factors are also associated with transcribing RNAPII on highly active lampbrush chromosomes (Gall et al. 1999).

3.7 Conformational changes in the CTD — connecting things

Interaction of the CTD with protein complexes involved in transcription and mRNA processing likely occurs in a highly dynamic and spatially regulated manner during the transcription cycle. How these processes are coordinated is still open to investigation, but recent genetic and biochemical evidence points to an intriguing and important role for peptidylprolyl-cis-trans-isomerases (PPIase) in the regulation of CTD function. PPIases catalyze cis-trans isomerization of the peptide bond on the amino-terminal sides of prolines in peptides and proteins.

Three distinct families of PPIases have been distinguished (Hunter 1998; Gothel and Marahiel 1999): the cyclophilins are sensitive to the immunosuppressant drug cyclosporin A; sensitivity to another immunosuppressant FK506 defines the second family, called the FK506 binding protein (FKBP) family (Hunter 1998; Gothel and Marahiel 1999); members of the third distinct family, the parvulin family, are not sensitive to these drugs. PPIases are thought to participate in protein folding, trafficking, assembly, and disassembly and regulation of protein activity (Hunter 1998; Gothel and Marahiel 1999). SRcyp/SCAF10, a member of the cyclophilin family of PPIases (Bourquin et al. 1997), and Ess1p/hPIN1 (Morris et al. 1999), a member of the parvulin family, can bind phosphorylated CTD (see also Table 6). I discussed SRcyp/SCAF10 above, and it is possible that proline isomerization of the CTD by SRcyp/SCAF10
determines the association of splicing components with the hyperphosphorylated CTD. Because Ess1p function might be connected to Fcp1p, I shall discuss the roles of Ess1p in more detail below.

Ess1p was the first eukaryotic parvulin to be discovered (Hanes et al. 1989). The ESS1 gene is essential for viability in yeast and the protein is highly conserved among eukaryotes (Hanes et al. 1989). Ess1p contains an amino-terminal WW domain and a carboxy-terminal PPIase domain. WW domains are protein-protein interaction domains that center around two adjacent, conserved Trp residues in a hydrophobic pocket that binds Pro-rich ligands. The WW domain of the human homolog of Ess1p, PIN1, binds peptides in which phosphorylated Ser or Thr residues are followed by Pro (Lu et al. 1999). The isomerization of prolines is dramatically enhanced about 1000-3000 fold when the proline residue is preceeded by either phosho-Ser or phosho-Thr rather than the same residue in its unphosphorylated form (Yaffe et al. 1997; Hani et al. 1999). The phosphorylated CTD therefore potentially provides two ideally positioned Pro residues in one YSPTSPS heptapeptide repeat and this has been confirmed by a structural analysis of a PIN1- phospho-CTD peptide complex (Verdecia et al. 2000).

Although initially identified as a protein involved in cell-cycle progression (Hanes et al. 1989; Lu et al. 1996; Shen et al. 1998; Winkler et al. 2000), several lines of evidence point to a major role for Ess1p/PIN1 in mRNA synthesis. Yeast strains with mutated ess1 have decreased mRNA levels for several genes (Wu et al. 2000). Affinity chromatography experiments using a CTD preparation that was phosphorylated by CTDK-I show that Ess1p is the major CTD-binding protein in yeast extracts (Morris et
al. 1999). Importantly, the interaction is dependent on phosphorylation of the CTD and is mediated by the WW domain of Ess1p (Morris et al. 1999). It was not tested, however, whether Ess1p can also bind to CTD phosphorylated by other CTD kinases such as TFIIH or Srb10p/Srb11p (Morris et al. 1999). However, RNAPII also was found to be the major PIN1-interacting protein in human interphase cells (Albert et al. 1999). This interactions likely reflect the abundance of phosphoSer-Pro dipeptides in the phosphorylated CTD.

Mutated alleles of ess1 were independently identified in a genetic screen that was designed to be very specific for mRNA 3’end formation in yeast (Hani et al. 1995; Hani et al. 1999). Yeast cells carrying these mutant alleles accumulate mRNAs with unprocessed 3’ ends at the non-permissive temperature (Hani et al. 1995; Hani et al. 1999). Thus, these results suggest that Ess1p/PIN1 functions in mRNA 3’end formation by linking the processing reaction to transcription via the hyperphosphorylated CTD of RNAPII.

A collection of ess1 mutants with changed amino acids in either the WW domain or the PPIase domain was used to identify genes that could suppress the temperature dependent lethality of the ess1 strains when present in high copy number (Wu et al. 2000). Interestingly, none of the suppressors encode cell cycle regulators. Instead, the suppressing genes encode proteins that bind DNA or modify chromatin structure or are regulatory factors or subunits of RNAPII (Wu et al. 2000). Among the suppressor genes is a TFIIS-like gene from S. cerevisiae and the gene encoding the RNAPII Rpb7p subunit from Candida albicans (Wu et al. 2000). ESI1 also genetically interacts with RPBI,
which encodes the largest subunit of RNAPII, and *SRB2*. Another interesting transcription-related suppressor is *CTHI*, which encodes a Zn\(^{2+}\)-finger protein (Wu et al. 2000). The human homologue of *CTHI*, TIS11, can bind to AU-rich sequences in the 3' ends of mRNAs and promote mRNA degradation (Carballo et al. 1998). Overexpression of TIS11, however, can stabilize certain mRNAs (Lai et al. 1999), and hence it was speculated that overexpression of *CTHI* might compensate for defects in 3' end formation caused by altered versions of Ess1p (Wu et al. 2000).

The *FCP1* gene can also suppress the temperature-sensitivity of *ess1* strains when present on a high-copy number plasmid (Wu et al. 2000). This suppression is dependent on intact Fcp1p, as the mutant alleles *fcpl-1* and *fcpl-2*, which are described in Chapter 2 of this thesis, are unable to rescue the lethality of *ess1* mutant strains (Wu et al. 2000). These results suggest that isomerization of the CTD by Ess1p might promote CTD dephosphorylation by Fcp1p, and further suggests that dephosphorylation of the CTD may be important for 3' end formation. Compromised Ess1p function could then be compensated by high concentrations of Fcp1p. It was not demonstrated, however, that *ess1* strains have an abnormal CTD phosphorylation pattern number (Wu et al. 2000).

3.8 Physiological conditions that alter phosphorylation of the RNAPII CTD

3.8.1 Diauxic shift

Yeast cells undergo a metabolic transition late in their exponential growth phase in glucose-based medium when they shift from fermentation to respiratory metabolism. During this diauxic shift, the CTD transiently accumulates phosphorylation on Ser2,
whereas Ser5 phosphorylation does not change (Patturajan et al. 1998a). This phospho-
Ser2 accumulation is not found in yeast cells with a deletion of ctk1, encoding the kinase
subunit of the CTD kinase CTDK-I, and this strain also has defects in the activation of
certain genes that normally occurs during diauxic shift, such as those encoding glycogen
synthase and cytosolic catalase (Patturajan et al. 1999). The amount of another CTD
kinase, Srb10p, is strongly reduced when yeast enter stationary phase, although it is not
known how this relates to CTD phosphorylation (Holstege et al. 1998)

3.8.2 Heat Shock

Heat shock causes an increase in Ser2 phosphorylation and at the same time a
decrease in the hypophosphorylated Rpb1p in *S. cerevisiae* (Patturajan et al. 1998a). The
increased phosphorylation of Ser2 still occurs in a ctk1Δ strain, indicating that CTDK-I is
not required for this heat shock response in *S. cerevisiae* (Patturajan et al. 1999).

Heat shock of HeLa cells also causes a change in the phosphorylation pattern of
RNAPII. Mild heat shock results in the disappearance of hyperphosphorylated
RNAPIIO but under more severe conditions, this form is predominant (Venetianer et al.
1995), likely as a result of the CTD kinase activity of the MAP kinases ERK1/ERK2
(Dubois et al. 1997). Heat shock also causes inhibition of the CTD kinase activity of
TFIIH (Dubois et al. 1997) as well as the CTD phosphatase activity of FCP1 (Dubois et
al. 1999) by creating forms of these enzymes that are insoluble. These changes in HeLa
cells coincide with the disappearance of the CC3 epitope that predominantly recognizes
phospho-Ser2 of the CTD heptapeptide repeat (Dubois et al. 1997).

3.8.3 Cell Cycle Regulation of Transcription and the RNAPII CTD

A number of different studies have shown that regulation of CTD phosphorylation might be involved in positive and negative regulation of the cell cycle. Cellular mitosis is accompanied by a loss of transcriptional activity by all three nuclear RNAPs and distinct changes in the CTD of RNAPII. Since hypophosphorylated RNAPIIA virtually disappears and hyperphosphorylated RNAPIIO predominates (Shermoen and O'Farrell 1991; Parsons and Spencer 1997), mitotic RNAPII is probably unable to initiate transcription. The p34CDC2 cyclin-dependent kinase subunit of the mitosis promoting factor (MPF; cdc2/cyclinB) can phosphorylate RNAPII, and this leads to inhibition of transcription in vitro (Gebara et al. 1997). However, it is also possible that inactivation of RNAPII activity during mitosis is not due to direct phosphorylation of the CTD by MPF (e.g. there may be a downstream cascade of kinases). Additionally, MPF phosphorylates the CDK7, p36 and p62 subunits of TFIIF during mitosis, leading to inhibition of the CTD kinase and transcription activities of TFIIF (Akoulitchev and Reinberg 1998; Long et al. 1998).

At the G1/S boundary of the cell cycle, regulation of the CTD kinase activity of TFIIF by the tumor suppressor p16INK4A may contribute to cell cycle regulation. P16INK4A can form a complex with TFIIF and RNAPII in vitro and might co-localize with CDK7 in vivo (Serizawa 1998). Importantly, p16INK4A specifically inhibits the CTD kinase activity of TFIIF but not the CDK2 kinase activity of TFIIF and this contributes
to the ability of p16<sup>INK4A</sup> to induce cell cycle arrest (Nishiwaki et al. 2000).

Treatment of human peripheral blood mononuclear cells with the mitogens phorbol myristate acetate or phytohemagglutinin leads to activation of these cells and a general increase in S-phase (Garriga et al. 1998). Interestingly, CyclinT1 and Cdk9 levels also increase, and the resulting increase in CTD kinase activity of P-TEFb promotes HIV replication by increasing the transcription of the HIV genome (Garriga et al. 1998). Serum stimulation of quiescent cells, which can cause G<sub>0</sub> cells to re-enter the cell cycle, leads to CTD phosphorylation by the MAP kinase ERK1/ERK2 (Dubois et al. 1994). but the role of these phosphorylation events is unclear.

3.8.4 The germ-cell lineage and oocyte fertilization

In early <i>Caenorhabditis elegans</i> embryos production of new mRNAs is inhibited in the germ line blastomeres. This inhibition is dependent on the presence of the germline factor PIE-1 and correlates with the absence in germline blastomeres of the H5 phosphoepitope, corresponding to phosph-Ser2, on the CTD (Seydoux et al. 1996; Seydoux and Dunn 1997). PIE-1 can repress transcription in cell culture assays when fused to the Gal4 DNA-binding domain (Batchelder et al. 1999). This repression involves a motif in PIE-1 whose primary sequence YAPMAPT somewhat resembles the CTD heptapeptide repeat (Batchelder et al. 1999), suggesting that the repression motif might target a protein complex that can bind the CTD and phosphorylate Ser2.

Mature <i>Xenopus</i> oocytes have a cell cycle arrest in meiotic metaphase II and are transcriptionally inactive (Bellier et al. 1997b). The cell cycle arrest and shutdown in
transcription coincide with the appearance of a large amount of hyperphosphorylated RNAPII as a consequence of the CTD kinase activity of the MAP kinase ERK1/ERK2 (Bellier et al. 1997b). RNAPII remains inactive for the first few cell divisions after fertilization (Bellier et al. 1997a). Resumption of transcription at later stages of zygote development in the rabbit and mouse is characterized by the re-appearance of RNAPIIA, inactivation of ERK1/ERK2 and translocation into the nucleus of RNAPII, which is largely outside the pronucleus at fertilization (Bellier et al. 1997a).

3.8.5 Genotoxic Agents

Exposure of cells to ionizing radiation activates the tyrosine kinase c-Abl by inducing the protein kinase ATM to phosphorylate a regulatory serine at position 465 of c-Abl (Baskaran et al. 1997b). This results in a fourfold increase in the tyrosine phosphorylation of the RNAPII CTD (Baskaran et al. 1997b). Similarly, the DNA damaging agent methanosulfonic methyl ester (MMS) causes an increase in tyrosine phosphorylation of the CTD (Liu et al. 1996). C-Abl is not required for the cell cycle arrest induced by these genotoxic agents (Liu et al. 1996), but the increased CTD phosphorylation by c-Abl on tyrosine might contribute to altered expression of certain genes after DNA damage (Baskaran et al. 1999).

Exposure of cells to UV radiation causes cell cycle arrest to provide time for the cell to repair the DNA damage before the next S or M phase. In mammalian cells, this occurs in part due to the upregulation of the cellular guardian p53. P53 interacts with TFIIH (Xiao et al. 1994; Leveillard et al. 1996) and wild-type but not an altered form of
p53 can inhibit the RNAPII CTD kinase activity of TFIIH in vitro (Schneider et al. 1998). Perhaps as a consequence, exposure of mouse embryonic fibroblasts to UV leads to a decrease of TFIIH CTD kinase activity (Adamczewski et al. 1996), likely in a p53 dependent manner (Schneider et al. 1998).

4. Ubiquitination of RNAPII

The ubiquitin-dependent proteolysis system is characterized by the covalent ligation of a multi-ubiquitin chain to target proteins (Hershko et al. 2000; Varshavsky et al. 2000). Ubiquitin is a 76 amino acid peptide which gets linked via its terminal glycine to a cysteine residue of the E1 ubiquitin activating enzyme. Besides the E1 enzymes, the E2 ubiquitin-conjugating enzyme and the E3 ubiquitin ligases are involved in the ligation of ubiquitin to protein substrates. The attachment of multiple ubiquitin moieties to proteins serves as a signal for recognition and degradation of these proteins by the 26S proteasome. The E3 proteins are thought to have a major role in determining the substrate specificity of the system.

4.1. Rsp5p ubiquitin ligase and the CTD

The S. cerevisiae RSP5 gene encodes an essential E3 enzyme that is a member of the HECT (homologous to E6-AP carboxy terminus) domain family of ubiquitin ligases. RSP5 was identified in several different genetic screens, and several cellular roles for Rsp5p have been suggested. Rsp5p contains an amino-terminal CS domain involved in binding membrane phospholipids, three WW domains, and a HECT domain that
constitutes the E3 ubiquitin ligase activity.

The Rbp1p subunit of yeast RNAPII is a substrate for Rsp5p. RNAPII can associate with Rsp5p in yeast extracts and reduced RSP5 expression leads to an elevated steady state level of Rpb1p (Huibregtse et al. 1997). The CTD of Rpb1p is necessary and sufficient for binding Rsp5p (Huibregtse et al. 1997). Conversely, the second, and possibly the third, WW domain of Rsp5p mediates the binding of Rsp5p to the CTD (Wang et al. 1999; Chang et al. 2000). Phosphorylation of serine, threonine and tyrosine residues within a minimal CTD sequence that is sufficient to strongly bind Rsp5p decreases the binding between the WW domains of Rsp5p and the CTD (Chang et al. 2000). Therefore, unphosphorylated or partly phosphorylated CTD might be a signal for an interaction between the CTD and Rsp5p ubiquitin ligase, targeting RNAPII for degradation by the proteasome.

4.2. Ubiquitination and transcription

The largest subunit of RNAPII is ubiquitinated during transcription reactions in vitro and this depends on phosphorylation of the CTD. Therefore, phosphorylation of the CTD might signal degradation of RNAPII during the regulation of transcription and DNA repair (Mitsui and Sharp 1999). Nevertheless, Rsp5p binds preferentially to hypophosphorylated CTD (Chang et al. 2000) and it is possible that another ubiquitin ligase is responsible for the preferential ubiquitination of hyperphosphorylated RNAPII during transcription.
4.3 UV radiation and DNA repair

UV radiation and cisplatin, which create bulky, helix-distorting lesions in the DNA, induce the ubiquitination of the largest subunit of RNAPII. Evidence suggests that the hyperphosphorylated RNAPII form is preferentially ubiquitinated in mammalian cells (Bregman et al. 1996). The ubiquitinated form of RNAPII is degraded through the proteosomal pathway and therefore, the steady state level of RNAPII is reduced dramatically after UV radiation (Ratner et al. 1998). Despite the finding that the hyperphosphorylated form of RNAPII is preferentially ubiquitinated, the level of the hypophosphorylated form is reduced as well (Ratner et al. 1998). After successful repair of the DNA damage, the levels of both forms of RNAPII eventually return to baseline. Interestingly, the ubiquitination of the largest subunit of RNAPII does not occur in fibroblasts from Cockayne’s syndrome patients exposed to UV radiation (Bregman et al. 1996).

UV irradiation and treatment of yeast cells with 4-NQO induce Rsp5p mediated ubiquitination and degradation of Rpb1p in vivo (Beaudenon et al. 1999). Surprisingly, transcription-coupled repair is not compromised in a rbp5 mutant strain, demonstrating that ubiquitination of Rpb1p is not required for transcription-coupled repair (Lommel et al. 2000). Hence, these studies suggest that failure to ubiquitinate the largest subunit of RNAPII in CS cells does not cause their defect in transcription-coupled repair. A revised model was therefore proposed in which elongating RNAPII stalls when it encounters a DNA lesion in normal cells leading to removal RNAPII from the chromatin followed by ubiquitination and degradation (Lommel et al. 2000). Cells from CS patients would have a
defect in their ability to remove the stalled elongation complex from the chromatin, and hence there would not be any free RNAPII largest subunit that can be ubiquitinated. It has not been determined whether ubiquitination of RNAPII takes place in the nucleus or the cytoplasm. The human homologue of Rsp5p, NEDD4 has been shown to be exclusively cytoplasmic (Anan et al. 1998), but the localization of Rsp5p in yeast is not known. If ubiquitination and degradation of RNAPII indeed occur in the cytoplasm, there must be a signaling mechanism that leads to nuclear export of RNAPII after a cell encounters genotoxic stress.

5. Thesis Outline

The work presented in this thesis is the initial characterization of the Fcp1p CTD phosphatase in the yeast *S. cerevisiae*. I will describe in Chapter 2 that Fcp1p is an unusual eukaryotic protein phosphatase that is required for dephosphorylation of the CTD and for most mRNA synthesis in yeast. As described in this introduction, CTD phosphatase is regulated by the RAP74 subunit of TFIIF and TFIIB. Chapter 3 contains a detailed examination of these protein-protein interactions and also evidence for the importance of the RAP74-Fcp1p interaction for Fcp1p function *in vivo*. In Chapter 4, I present evidence for an opposing role of Fcp1p and the CTD kinase Srb10p in the regulation of CTD phosphorylation and also describe new phenotypes for yeast cells with mutations in *fcp1*. Reduced activity of Fcp1p causes increased synthesis of the largest subunit of RNAPII that is not assembled into RNAPII complexes. This phenomena is described in Chapter 5 of this thesis. Finally, I summarize the findings
presented in this thesis and suggest further directions in Chapter 6.
Chapter 2

A New Type of Eukaryotic Protein Phosphatase Required for Dephosphorylation of the CTD and Transcription by RNA polymerase II in *Saccharomyces cerevisiae*

A version of this Chapter has been published in “Molecular Cell”, Vol 4, pp 55-62 (1999)

I made all the DNA constructs and yeast strains that were used in this study. I performed the experiments presented in Fig. 1, Fig. 3, Fig. 4A and 4B and Fig. 5B. For the experiments in Fig. 3, I was assisted by Fiona Kouyoumdjian, a summer student working with me, and David Jansma, a graduate student in Jim Friesen’s laboratory. The experiments in Fig. 2 were done by Bill Lester, the experiments in Fig. 4C and Fig. 5A were performed by Opher Gileadi (The Weizman Institute, Israel). The genome wide expression data presented in Table 1 were obtained by Frank Holstege in Richard Young’s laboratory (Whitehead Institute for Biomedical Research, USA)
Abstract

The carboxy-terminal domain (CTD) of the largest subunit of RNA polymerase II is phosphorylated soon after transcriptional initiation. We show here that the essential *FCPI* gene of *Saccharomyces cerevisiae* is linked genetically to RNA polymerase II and encodes a CTD phosphatase essential for dephosphorylation of RNA polymerase II *in vivo*. *Fcp1p* contains a phosphatase motif, $\Psi \Psi \Psi D\Psi D\Psi (T/V)\Psi\Psi$, which is novel for eukaryotic protein phosphatases and essential for *Fcp1p* to function *in vivo*. This motif is also required for recombinant *Fcp1p* to dephosphorylate the RNA polymerase II CTD or the artificial substrate p-nitrophenylphosphate *in vitro*. The effects of *fcp1* mutations in global run-on and genome-wide expression studies show that transcription by RNA polymerase II in *S. cerevisiae* generally requires CTD phosphatase.
Introduction

Transcriptional initiation by RNA polymerase II (RNAPII) in *Saccharomyces cerevisiae* generally utilizes a preassembled RNAPII holoenzyme containing both general transcription factors and SRB/mediator proteins that regulate activation or repression of transcription (Myer and Young 1998). The largest subunit of RNAPII contains a C-terminal domain (CTD) consisting of tandem repeats with the consensus sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser. Mammalian RNAPII has 52 repeats, whereas *S. cerevisiae* RNAPII has 26 or 27 repeats. The CTD is essential for yeast growth and for association of the mediator with RNAPII but its role in mRNA synthesis is not completely understood. There is evidence that the CTD functions during initiation, promoter clearance, and chain elongation by RNAPII and also that it has an important role in transcript processing (Neugebauer and Roth 1997).

Whereas RNAPII molecules in the transcription initiation apparatus have CTDs that lack phosphate, the CTD becomes heavily phosphorylated during or shortly after transcription initiation and appears to remain so during transcript elongation *in vivo* (Cadena and Dahmus 1987; O'Brien et al. 1994). The hyperphosphorylated form of RNAPII is designated RNAPIIO, whereas the nonphosphorylated form is designated RNAPIIA. Several cyclin-dependent kinases have been implicated in the phosphorylation of the CTD, including the yeast TFIIH subunit Kin28p and its human homologue CDK7 (Feaver et al. 1994; Valay et al. 1995), the *Drosophila* and human P-TEFb kinases (Marshall et al. 1996; Zhu et al. 1997; Peng et al. 1998b) and the yeast RNAPII holoenzyme component, Srb10p (Hengartner et al. 1998).
Several lines of evidence suggest that the II A form of RNAPII is recruited to promoters. It is predominantly this hypophosphorylated form of RNAPII that can initiate transcription in vitro (Lu et al. 1991; Chesnut et al. 1992). The SRB/mediator complex associates preferentially with RNAPIIA, and purified RNAPII holoenzyme, which is competent for promoter-dependent initiation, lacks phosphate on its CTD (Koleske and Young 1994; Svejstrup et al. 1997). If RNAPII A alone can initiate transcription in vivo, the hyperphosphorylated CTD in elongation complexes must be dephosphorylated in order to regenerate RNAPII A for subsequent rounds of transcription.

CTD phosphatase activities have been purified from HeLa cells and S. cerevisiae (Chambers and Dahmus 1994; Chambers and Kane 1996). The CTD phosphatase from HeLa cells is stimulated by the RAP74 subunit of the general transcription factor TFIIF, and this stimulation can be inhibited by TFIIB (Chambers et al. 1995). A two hybrid screen with RAP74 as bait led to the cloning of cDNAs encoding FCP1, a human protein that co-purifies and co-immunoprecipitates with CTD phosphatase activity (Archambault et al. 1998b). The homologous yeast FCP1 gene is an essential gene in S. cerevisiae, and there are two adjacent RAP74-binding regions toward the carboxy-terminus of yFcp1p (Archambault et al. 1997) (see Fig. 1A). Recombinant yeast Fcp1p could replace one of two fractions apparently required for yeast CTD phosphatase activity in vitro (Archambault et al. 1997).

At the time that we reported the sequences of human and yeast FCP1, we could not identify protein phosphatase motifs in FCP1 and so could not decide whether FCP1
was the catalytic or regulatory subunit of CTD phosphatase. In this study we show that yeast Fcp1p is a new type of eukaryotic protein phosphatase. Furthermore, we show that Fcp1p is required for dephosphorylation of the CTD and essential for most transcription by RNAPII \textit{in vivo}.
Results

**Fcp1p contains a functionally important phosphatase motif**

The absence of a known protein phosphatase motif in the amino acid sequence of Fcp1p, together with the apparent requirement for a second fraction, in addition to renatured Fcp1p, for yeast CTD phosphatase activity, made it difficult to know whether Fcp1p was the catalytic or regulatory subunit of CTD phosphatase (Archambault et al. 1997; Archambault et al. 1998b). Recently, however, we realized that the motif \( \Psi^+\Psi^+DXDX(T/V)\Psi^+\Psi^+ \), characteristic of a new family of small molecule phosphotransferases and phosphohydrolases (Collet et al. 1998), is also found in perfectly conserved form in the FCP-homology regions of Fcp1p and all but one database protein we had originally assigned to the FCP1 family (Fig. 1A).

The high degree of conservation of the \( \Psi^+\Psi^+DXDX(T/V)\Psi^+\Psi^+ \) motif and surrounding amino acids in Fcp1p and related proteins suggested that this motif had functional significance. Indeed, the first aspartate of this motif has been shown to act as a phosphoryl acceptor during catalysis by human phosphomannomutase, and mutagenesis of human L-3-phosphoserine phosphatase also implicated this first aspartate as the residue that becomes phosphorylated during catalysis (Collet et al. 1998). These results, in addition to the fact that several of the enzymes listed in Fig. 1 have been shown to act via a phosphoenzyme intermediate, led to the suggestion that the \( \Psi^+\Psi^+DXDX(T/V)\Psi^+\Psi^+ \) motif is a conserved phosphoryl acceptor site in this family of enzymes.
Figure 1. Fcp1p contains a sequence motif essential for it to function in vivo that is found in a family of small molecule phosphotransferases and phosphohydrolases. (A) Alignment of the most conserved part of the FCP-homology domain with similar regions in a family of small molecule phosphotransferases and phosphohydrolases. Shown in the upper part of the panel is a schematic diagram indicating the FCP-homology, BRCT and RAP74-binding domains of yeast Fcp1p. The most conserved portions of the FCP-homology domain are highlighted in black, and one of these motifs is aligned with amino acid sequences in other members of the FCP family and the indicated small molecule phosphatases and phosphotransferases (Collet et al. 1998).

(B) The conserved aspartate residues in the $\Psi^\Psi^\Psi DXDX(T/V)\Psi^\Psi$ phosphatase motif are essential for Fcp1p function in vivo. Strains with a wild type copy of FCP1 on the URA3 CEN/ARS plasmid pRS316 (Sikorski and Hieter 1989) and a chromosomal FCP1 deletion (fcp1Δ:: LEU2) were transformed with TRPI CEN/ARS plasmids (pRS314) carrying the wild type FCP1 gene, no FCP1 gene, or fcp1 genes with the indicated mutations in its $\Psi^\Psi^\Psi DXDX(T/V)\Psi^\Psi$ motif. The abilities of these mutant alleles to complement the FCP1 chromosomal deletion were tested by plasmid shuffling on synthetic complete plates lacking Leu and Trp and containing 5-FOA to counterselect the URA3 marker.
**A**

**FCP Hom.**

**BRCT R74R74**

<table>
<thead>
<tr>
<th>Species</th>
<th>Accession</th>
<th>Sequence</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>H. sapiens</em></td>
<td>AF081237</td>
<td>LVVLVTTNT</td>
<td>P-Ser-Hpr phosphatase</td>
</tr>
<tr>
<td><em>S. cerevisiae</em></td>
<td>Q03254</td>
<td>LVVLVTTNT</td>
<td>Phosphoserine phosphatase</td>
</tr>
<tr>
<td><em>S. cerevisiae</em></td>
<td>S64841</td>
<td>LVVLVTTNT</td>
<td>Phosphoglycerate phosphatase</td>
</tr>
<tr>
<td><em>S. cerevisiae</em></td>
<td>P38757</td>
<td>LVVLVTTNT</td>
<td>Deoxyglucose 6-P phosphatase</td>
</tr>
<tr>
<td><em>S. cerevisiae</em></td>
<td>S64752</td>
<td>LVVLVTTNT</td>
<td>Glycerol 3-P phosphatase</td>
</tr>
<tr>
<td><em>S. pombe</em></td>
<td>Q09695</td>
<td>LVVLVTTNT</td>
<td>Histidinol phosphatase</td>
</tr>
<tr>
<td><em>B. campestris</em></td>
<td>L46538</td>
<td>LVVLVTTNT</td>
<td>Trehalose 6-P phosphatase</td>
</tr>
<tr>
<td><em>C. elegans</em></td>
<td>U29336</td>
<td>LVVLVTTNT</td>
<td>Paranitrophenyl phosphatase</td>
</tr>
<tr>
<td><em>M. musculus</em></td>
<td>W29399</td>
<td>LVVLVTTNT</td>
<td>Phosphomannomutase</td>
</tr>
<tr>
<td><em>H. sapiens</em></td>
<td>H24417</td>
<td>LVVLVTTNT</td>
<td>β-phosphoglucomutase</td>
</tr>
<tr>
<td><em>H. sapiens</em> OS-4</td>
<td>AF090152</td>
<td>LVVLVTTNT</td>
<td></td>
</tr>
<tr>
<td><em>H. sapiens</em> HYA22</td>
<td>D88153</td>
<td>LVVLVTTNT</td>
<td></td>
</tr>
<tr>
<td><em>B. subtilis</em></td>
<td>2618860</td>
<td>LVVLVTTNT</td>
<td></td>
</tr>
<tr>
<td><em>S. cerevisiae</em></td>
<td>P42341</td>
<td>LVVLVTTNT</td>
<td></td>
</tr>
<tr>
<td><em>H. sapiens</em></td>
<td>Y18275</td>
<td>LVVLVTTNT</td>
<td></td>
</tr>
<tr>
<td>E. coli</td>
<td>P32662</td>
<td>LVVLVTTNT</td>
<td></td>
</tr>
<tr>
<td>R. sphaeroides</td>
<td>1825472</td>
<td>LVVLVTTNT</td>
<td></td>
</tr>
<tr>
<td>E. coli</td>
<td>P77625</td>
<td>LVVLVTTNT</td>
<td></td>
</tr>
<tr>
<td><em>S. cerevisiae</em></td>
<td>P38774</td>
<td>LVVLVTTNT</td>
<td></td>
</tr>
<tr>
<td><em>S. cerevisiae</em></td>
<td>P41277</td>
<td>LVVLVTTNT</td>
<td></td>
</tr>
<tr>
<td><em>S. cerevisiae</em></td>
<td>P40106</td>
<td>LVVLVTTNT</td>
<td></td>
</tr>
<tr>
<td>E. coli</td>
<td>P06987</td>
<td>LVVLVTTNT</td>
<td></td>
</tr>
<tr>
<td>H. influenzae</td>
<td>P44327</td>
<td>LVVLVTTNT</td>
<td></td>
</tr>
<tr>
<td>E. coli</td>
<td>P31678</td>
<td>LVVLVTTNT</td>
<td></td>
</tr>
<tr>
<td><em>S. cerevisiae</em></td>
<td>J67669</td>
<td>LVVLVTTNT</td>
<td></td>
</tr>
<tr>
<td><em>C. elegans</em></td>
<td>P34492</td>
<td>LVVLVTTNT</td>
<td></td>
</tr>
<tr>
<td><em>S. cerevisiae</em></td>
<td>P07283</td>
<td>LVVLVTTNT</td>
<td></td>
</tr>
<tr>
<td><em>H. sapiens</em></td>
<td>Q92871</td>
<td>LVVLVTTNT</td>
<td></td>
</tr>
<tr>
<td>B. subtilis</td>
<td>O60956</td>
<td>LVVLVTTNT</td>
<td></td>
</tr>
<tr>
<td>L. lactis</td>
<td>P71447</td>
<td>LVVLVTTNT</td>
<td></td>
</tr>
</tbody>
</table>

**Phosphatase Motif:** $\psi \psi DXDX(T/Y)\psi$
To investigate the importance of the two conserved aspartates, D180 and D182, within this motif of yeast Fcp1p, we used site-directed mutagenesis and a plasmid shuffling procedure. Strains were created that contained both the wild type yeast FCP1 gene on a plasmid with a URA3 selective marker and either wild type FCP1 or various fcp1 mutant genes on plasmids containing a TRP1 selective marker. The plasmid with the URA3 marker was then counterselected on medium containing 5-fluoroorotic acid (5-FOA). Wild type FCP1, but none of the four fcp1 mutant alleles (D180E, D180N, D182E and D182N), complemented a chromosomal deletion of FCP1, although all four mutant alleles could be rescued by wild type FCP1 (Fig. 1B). Western blot analysis of protein extracts derived from strains containing both the fcp1-3 allele, which encodes an Fcp1p that lacks amino acids 2-134 but is able to maintain cell viability, and the fcp1 mutant alleles encoding full length proteins showed that the wild type and mutant proteins are expressed at similar levels (data not shown). The inability of the fcp1 alleles carrying mutations in the aspartate residues of the phosphatase motif to complement a chromosomal deletion of FCP1 suggested that Fcp1p functions as a phosphatase to support the survival of S. cerevisiae.

Fcp1p is a CTD phosphatase

The importance for yeast growth of the novel phosphatase motif in Fcp1p suggested that Fcp1p is a phosphatase. In order to test Fcp1p directly for phosphatase activity in vitro, we expressed recombinant His6-tagged Fcp1p in insect cells using a baculovirus-based expression system. The recombinant protein was then purified to near
homogeneity by nickel chelate affinity chromatography followed by ion exchange chromatography on a MonoQ column (Fig. 2A). As shown in Fig. 2A, highly purified recombinant yeast Fcp1p co-purified precisely with an activity that dephosphorylated the artificial phosphatase substrate p-nitrophenylphosphate, producing p-nitrophenol which absorbs strongly at 410 nm. By varying the substrate concentration, we estimated the $K_m$ to be approximately 60 mM, which is 50-100 times higher than those of alkaline phosphatases when assayed on the same substrate, and the $k_{cat}$ to be 2 sec$^{-1}$. The specific activity of highly purified wild type Fcp1p is 20 000 units/µg (See Material and Methods for unit definition). Importantly, although careful titration revealed that recombinant Fcp1p with the D182E mutation still retained 42% of the wild type activity, recombinant Fcp1p with the D180E mutation was completely inactive (Fig. 2A and data not shown). Since there was virtually no other protein present in the reaction, the ability of recombinant wild type Fcp1p and inability of recombinant Fcp1p with the D180E mutation to dephosphorylate p-nitrophenylphosphate demonstrated that Fcp1p is a phosphatase. This conclusion was confirmed by showing that recombinant Fcp1p that had been subjected to SDS-PAGE and renatured following its recovery from a gel slice was still able to hydrolyze p-nitrophenylphosphate (data not shown). The ability of Fcp1p to dephosphorylate p-nitrophenylphosphate eliminated the possibility that the Fcp1p, RNAPII or TFIIH that were present in all previous assays (Archambault et al. 1997; Archambault et al. 1998b; Marshall et al. 1998) of yeast Fcp1p or human FCP1 contained subunits or contaminants with phosphatase activity.
Figure 2. Recombinant Fcp1p can hydrolyze p-nitrophenylphosphate and dephosphorylate the CTD

(A) Wild type Fcp1p, as well as Fcp1p with amino acid changes D180E and D182E, were produced in the baculovirus system, purified and tested for the ability to hydrolyze p-nitrophenylphosphate. Assays and SDS-PAGE analysis were done on fractions eluted from a MonoQ column used as a final step to purify wild type and mutant Fcp1p. One enzyme unit produces 1 pmol nitrophenol per hour.

(B) The indicated amounts of recombinant wild type, D180E, and D182E Fcp1p proteins were incubated with 150 ng of ^{32}P-labelled RNAPII that had been phosphorylated by TFIIH. Dephosphorylation of the Rpb1p CTD was monitored by SDS-PAGE and quantitated by Phosphor Imager analysis. The control reaction contained no Fcp1p.
As shown in Fig. 2B, recombinant Fcp1p was also able to dephosphorylate RNAPII in the absence of any other factor. In these assays, RNAPII was phosphorylated on the CTD in reactions containing the general transcription factor TFIIH and [γ-32P]-ATP. Fcp1p acted as a CTD phosphatase as judged by the disappearance of radioactive label from RNAPII. This effect was dependent on the concentration of Fcp1p in the reaction, and 90% dephosphorylation was achieved at approximately equimolar amounts of Fcp1p and RNAPII. Most importantly, neither the D180E nor the D182E mutant protein had any activity in this assay (Fig. 2B). This was consistent with our observation that both mutations are lethal (Fig. 1B) and strongly suggested that Fcp1p is a CTD phosphatase.

**Genetic Interaction of FCP1 with the gene encoding the largest subunit of RNAPII**

In order to examine the role of Fcp1p in vivo, two conditional-lethal mutations were made in FCP1 by altering conserved residues in its broader FCP-homology region. Strains with the fcp1-l (R250A, P251A) mutation located carboxy-terminal to the \( \Psi\Psi DXDX(T/V)\Psi\Psi\) phosphatase motif and the fcp1-2 (L177A, L181A, H187A) mutation in less critical amino acids of the phosphatase motif were both viable at 30°C. Neither strain showed a cold-sensitive phenotype, but the fcp1-l strain was not viable at temperatures above 33°C, and the fcp1-2 strain had a growth defect at 37°C (Fig. 3A). Western blotting on extracts derived from these strains after shift to the non-permissive temperature revealed that fcp1-l strains have an Fcp1p that is degraded after shift to
Figure 3. Temperature-sensitive mutations in *FCPI* and genetic interaction with RNAPII

(A) The *fcp1*-1 and *fcp1*-2 mutations cause temperature-sensitive growth.

Strains with wild type and mutant *FCPI* genes on *TRP1* *CEN/ARS* plasmids in an *fcp1Δ::LEU2* chromosomal background were grown at the indicated temperatures on synthetic complete plates lacking Leu and Trp.

(B) Western blot analysis of yeast extracts prepared from *FCPI* wild type and *fcp1*-1 and *fcp1*-2 strains with an affinity-purified polyclonal antibody against Fcp1p. Fcp1p is degraded after the *fcp1*-1 mutant is shifted to 37°C. Western blotting with antibodies against the Tfb1p subunit of TFIH served as a loading control.

(C) *fcp1* mutations are synthetically lethal with a decreased amount of the RNAPII largest subunit. *FCPI*, *fcp1*-1, *fcp1*-2 and *fcp1*-3 strains in which the *RPB1* promoter was replaced with the *LEU2* promoter were tested for growth on synthetic complete plates either lacking leucine (Rpb1p expressed at normal levels) or containing 2 mM Leu (Rpb1p repressed 5-10 fold) (Archambault et al. 1996). Plates were incubated at 30°C until colonies were clearly visible.
37°C, whereas Fcp1p in the \textit{fcp1-2} strain is more stable at 37°C (Fig. 3B). The \textit{fcp1-1} mutation may create an unstable protein, and the observed temperature-sensitivity might be due to degradation of this essential gene product.

As an initial test for the involvement of Fcp1p in mRNA synthesis, we examined \textit{fcp1} mutations for synthetic lethality when they were combined with decreased levels of Rpb1p, the largest subunit of RNAPII. The level of the largest subunit is reduced 5-10 fold when leucine is added to the growth medium in strains that produce Rpb1p under the control of the \textit{LEU2} promoter instead of the natural \textit{RPB1} promoter (Archambault et al. 1996). Such strains with wild type \textit{FCP1}, the \textit{fcp1-1} or \textit{fcp1-2} point mutations, or the \textit{fcp1-3} deletion of amino-terminal residues 2-134 grew at similar rates in the absence of leucine. When the level of Rpb1p was reduced by the addition of leucine to the growth medium, the \textit{FCP1} strain had a slow-growing phenotype, whereas strains with the \textit{fcp1-1} and \textit{fcp1-2} point mutations and the \textit{fcp1-3} deletion mutation were unable to grow (Fig. 3C). Thus, mutations in \textit{FCP1} are synthetically lethal with a reduced amount of the largest RNAPII subunit. The genetic interaction between these molecules is consistent with a positive role for Fcp1p in transcription by RNAPII.

\textbf{Fcp1p is required for dephosphorylation of the CTD in vivo}

In order to test directly whether Fcp1p is required \textit{in vivo} to dephosphorylate Rpb1p, the largest subunit of RNAPII, we examined the state of the RNAPII CTD in strains in which Fcp1p activity had been destroyed by induced proteolysis or by thermal inactivation of a temperature-sensitive mutant protein. To induce the proteolysis of
Fcp1p, we created an *fcp1Δ* yeast strain with an *ANB-URLF-FCPI* fusion gene encoding a Ubiquitin-Fcp1p fusion protein (Moqtaderi et al. 1996). Addition of Cu⁺⁺ to the growth medium leads to repression of the fusion gene and proteolysis of the fusion protein. Compensatory effects on the expression of the *RPB1* gene (Guialis et al. 1979) caused by the proteolysis of Fcp1p were circumvented by replacing the *RPB1* promoter with the *LEU2* promoter (see also Chapter 5 for a detailed description). The amount of Rpb1p in such a strain is essentially normal when leucine is omitted from the growth medium (Archambault et al. 1996). This strain stopped growing within 8 hr after the addition to the growth medium of 0.5 mM CuSO₄ (Fig. 4A). Western blotting revealed that Fcp1p decreased to about 10% of the original level after two hours and persisted at 5-7% of the original level from four to eight hours after the addition of CuSO₄ (Fig. 4B) (the residual Fcp1p was slightly smaller, presumably because of removal of its ubiquitin tag). Importantly, western blots with polyclonal antibodies specific for the phosphorylated CTD showed a large increase in the amount of the hyperphosphorylated IIo form of Rpb1p (Fig. 4B). The signal obtained with the monoclonal antibody 8WG16, which bound preferentially in these conditions to the hypophosphorylated CTD in RNAPIIA (Patturajan et al. 1998a), was greatly reduced upon the destruction of Fcp1p (Fig. 4B). This low level of RNAPIIA persisted until the eight hour time point, presumably because of the 5-7% of the Fcp1p that was not destroyed after addition of CuSO₄ to the growth medium. The amount of the TFIIH
Figure 4. Fcp1p is required in vivo for dephosphorylation of the RNAPII CTD

(A) Effect on growth of Cu\textsuperscript{2+}-induced proteolysis of Fcp1p. Strains harboring either a conditional allele of FCP1, which allows both the degradation of Ubiquitin-tagged Fcp1p and the repression of ANB-URLF-FCPI mRNA synthesis (Moqtaderi et al. 1996), or the control TRP\textsuperscript{+} plasmid pRS314, in addition to an RPB1 allele that has its endogenous promoter replaced with the heterologous LEU2 promoter, were cultured in synthetic complete medium lacking Trp and Leu. Growth was followed for both strains after the addition of 0.5 mM CuSO\textsubscript{4}.

(B) Protein extracts prepared at various times after the addition of 0.5 mM CuSO\textsubscript{4} to the Ubiquitin-Fcp1p strain described in (A) were western-blotted with polyclonal antibodies specific for the phosphorylated form of the largest subunit of RNAPII and the monoclonal antibody 8WG16, which, in these conditions (Patturajan et al. 1998a), predominantly recognizes the hypophosphorylated form of the largest subunit. The same extracts were also probed with affinity-purified polyclonal antibodies against Fcp1p and the Tfb1p subunit of the general transcription factor TFIIH. The percentages of Fcp1p remaining in the extracts prepared after the induction of proteolysis with CuSO\textsubscript{4} were estimated by comparison with various amounts of the control extract electrophoresed on the same gel.

(C) Extracts prepared from FCP1 and fcpl-2 strains at the indicated times after a shift from 26°C to 37°C were western-blotted and probed with monoclonal antibody 8WG16 under conditions (Valay et al. 1995) that assess the relative amounts of hyperphosphorylated and hypophosphorylated forms of the largest subunit of RNAPII.
subunit Tfb1p, used as a loading control, changed only modestly (Fig. 4B). This result demonstrated that Fcp1p is needed for the dephosphorylation of RNAPII in vivo.

A similar experiment was done with a strain containing the fcp1-2 temperature-sensitive mutation after a shift to the non-permissive temperature of 37°C (Fig. 4C). In this experiment, western blotting conditions were used that allowed the 8WG16 monoclonal antibody to recognize both forms of RNAPII (Valay et al. 1995). As was the case when Fcp1p was proteolyzed, there was again a clear shift towards the hyperphosphorylated II0 form of RNAPII when cells were shifted to the non-permissive temperature (Fig. 4C), indicating that Fcp1p is required to dephosphorylate RNAPII in vivo. Even at the permissive temperature of 26°C, the phosphorylation pattern appeared abnormal in this strain, with the largest RNAPII subunit having a range of mobilities in an SDS gel indicative of intermediate states of phosphorylation.

Fcp1p is needed for global transcription by RNA polymerase II in vivo

In view of our genetic evidence that Fcp1p is involved in transcription by RNAPII, as well as our observation that Fcp1p is required to dephosphorylate RNAPII in vivo, we next investigated whether Fcp1p is generally required for transcription in S. cerevisiae.

First, we used a global run-on assay to measure transcription in the nuclei of temperature-sensitive strains fcp1-1 and fcp1-2. Cells were grown at 26°C, transferred to 37°C for 45 min, and then permeabilized with sarcosyl. The incorporation of radioactivity
Figure 5. Fcp1p is generally required for transcription by RNAPII \textit{in vivo}

(A) Global transcription is significantly reduced in \textit{fcp1-1} and \textit{fcp1-2} mutant strains after a shift to the non-permissive temperature. Wild type and mutant cells were either maintained at the permissive temperature (26°C) or shifted to the non-permissive temperature (37°C) for 45 min, permeabilized with sarcosyl, and assayed for total incorporation of radioactivity from $[\alpha^{-32P}]$-UTP.

(B) Effects of inactivating Fcp1p on poly(A)$^+$ mRNA levels. Total cellular RNA was prepared from \textit{FCPL}, \textit{fcp1-1}, and \textit{fcp1-2} strains at various times after a shift to 37°C. A $^{32}$P-labeled oligo-dT probe was hybridized to 1 µg of total slot-blotted RNA.
from \([\alpha-^{32}\text{P}]-\text{UTP}\) measured total transcription by RNA polymerases I, II and III. At the permissive temperature, the wild type strain incorporated \(3.8 \times 10^5\) cpm, whereas the \(fcp1-1\) and \(fcp1-2\) mutant strains had somewhat reduced incorporation of \(2.5 \times 10^5\) cpm and \(1.5 \times 10^5\) cpm, respectively. After a 45 min shift to the nonpermissive temperature, run-on transcription was still \(2.8 \times 10^5\) cpm in the \(FCP1\) wild type strain, but was substantially reduced to \(0.4 \times 10^5\) cpm in the \(fcp1-1\) strain and \(0.7 \times 10^5\) cpm in the \(fcp1-2\) strain (Fig. 5A). Although this experiment did not distinguish among the three nuclear RNA polymerases, the decrease in run-on transcription caused by the \(fcp1\) mutations was similar to the decrease caused by the \(rpb1-1\) mutation in the largest subunit of RNAPII (O. Gileadi, unpublished data). Therefore, it seemed likely that the \(fcp1\) mutations had caused a substantial decrease in the amount of transcription by RNAPII.

In order to assess directly the role of Fcp1p in the transcription of mRNA-encoding genes, we measured the total amount of poly(A)\(^{+}\) mRNA by hybridizing total cellular RNA from wild type, \(fcp1-1\) and \(fcp1-2\) strains with a \(^{32}\text{P}\)-labeled oligo(dT) probe. In each case, the cellular RNAs were first analyzed by gel electrophoresis to ensure the quality and integrity of our RNA preparations, and the extracted RNA at every time point after a shift to 37\(^\circ\)C was adjusted to contain the same amount of 28S and 18S ribosomal RNAs (data not shown). Not surprisingly, the total amount of poly(A)\(^{+}\) mRNA did not change after the wild type strain was shifted to the non-permissive temperature (Fig. 5B). There was, however, a dramatic decline in the amount of poly(A)\(^{+}\) mRNA when both the \(fcp1\) mutant strains were shifted to 37\(^\circ\)C (Fig. 5B). This decline
was more pronounced in the *fcpl-l* strain than in the *fcpl-2* strain, consistent with the finding that the *fcpl-l* strain is more tightly temperature-sensitive than the *fcpl-2* strain (Fig. 3A). Based on these results, we concluded that Fcp1p must be required for the transcription of most genes in *S. cerevisiae*.

To confirm these results and identify all genes that require Fcp1p for transcription, we employed high-density oligonucleotide arrays to measure transcription from all yeast genes in a single experiment (Holstege et al. 1998). The genome-wide expression profile of the temperature-sensitive *fcpl-l* strain was compared in two independent experiments to an isogenic *FCPL* wild type strain 45 min after both strains had been shifted to 37°C. Of the 6179 yeast genes represented on the arrays, the transcript levels from 5864 genes were reproducibly determined and could be analyzed, and the results are summarized in Table 1. Of the 5846 genes whose transcripts were quantitated, the levels of 4502 mRNAs (77%) decreased by more than twofold in the *fcpl-l* mutant relative to wild type cells, whereas the levels of the remaining mRNAs were not significantly affected. Hence, the *fcpl-l* mutation causes a nearly global defect in transcription by RNAPII.

The reduction in the level of a particular mRNA observed in temperature-sensitive mutants soon after a shift to the non-permissive temperature could be a primary effect of the mutation or else a secondary effect caused by a substantial reduction in the amount of some other transcript and its translation product. To distinguish primary effects from those that are potentially secondary, we used an approach that has been developed to identify genes whose change is almost certainly a direct consequence of the loss of
Table 1. Dependence of Genome-Wide Transcription on Fcp1p

<table>
<thead>
<tr>
<th>Gene Set</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yeast genes represented on arrays</td>
<td>6179</td>
</tr>
<tr>
<td>Genes whose mRNA levels were determined</td>
<td>5846</td>
</tr>
<tr>
<td>Genes whose mRNA decreased &gt;2-fold</td>
<td>4502 (77%)</td>
</tr>
<tr>
<td>Genes appropriate for comparison withpbl-1</td>
<td>5240</td>
</tr>
<tr>
<td>Genes whose mRNA decay rates fit those ofpbl-1</td>
<td>3616 (69%)</td>
</tr>
</tbody>
</table>

The results of genome-wide transcription analysis following temperature-sensitive inactivation of Fcp1p in thefcp7-1 mutant strain are summarized. Approximately 6% of genes produce transcripts with half lives greater than 45 min, and in these cases, the requirement for Fcp1p has not been assessed. Detailed results can be found on the world wide web at http://www.wi.mit.edu/young/CTD_phosphatase.html
function of a temperature-sensitive factor (Holstege et al. 1998). In this approach, the
decay rates for all transcripts are compared with those determined for the RNAPII
temperature-sensitive mutant rpb1-1 under identical conditions. By identifying the
transcripts with similar mRNA decay rates in the rpb1-1 and fcp1-1 mutants, this method
identifies those genes whose expression is equivalently dependent on Fcp1p and RNAPII
itself.

Of the 5864 transcript levels scored in the fcp1-1 mutant, 5240 could be
accurately compared with transcript levels after inactivation of RNAPII. Of these 5240
genes whose mRNA levels declined twofold or more in the rpb1-1 strain, 3616 (69%) were reduced to a similar extent in both the fcp1-1 mutant and the RNAPII mutant.
Because we do not know how quickly Fcp1p is inactivated by the fcp1-1 mutation after a
shift to 37°C, we concluded that the expression of at least 69% of expressed yeast genes
is as dependent on Fcp1p as it is on RNAPII itself. Previous studies have shown that the
equivalent values are 93% for the RNAPII holoenzyme component Srb4p, 87% for the
TFI1H kinase subunit Kin28p, 57% for the TAF117 component of the TFIID and
SAGA complexes, and 54% for the large subunit of the general transcription factor TFIIE
(Apone et al. 1998; Holstege et al. 1998). Thus, the inactivation of Fcp1p leads to a
severe and nearly global defect in RNAPII-dependent transcription, suggesting that
Fcp1p is a positive transcription factor that is required for the synthesis of most
mRNAs.
Discussion

**Fcp1p is a CTD phosphatase**

Protein phosphatases have been classified into three structurally and functionally diverse families (Barford et al. 1998). We believe that Fcp1p could be the founding member of a new family of eukaryotic protein phosphatases. Fcp1p does not contain any amino acid sequence motifs which are characteristic of known eukaryotic protein phosphatase families, but does contain the motif ΨΨΨDXDX(T/V)ΨΨ which was first described for a family of small molecule phosphatases and phosphotransferases (Collet et al. 1998). In this family, the first aspartate in this motif participates in catalysis and is essential for activity, whereas the second aspartate varies in importance (Collet et al. 1998). We have shown here that both aspartates are essential for Fcp1p to function *in vivo* and for recombinant Fcp1p to dephosphorylate the CTD of RNAPII *in vitro*. Since only the first aspartate of Fcp1p is essential for activity with p-nitrophenylphosphate, the second aspartate is likely to have a role in substrate recognition rather than catalysis.

We had previously observed that a number of other database proteins of unknown function shared an FCP-homology region with yeast and human FCP1 (Archambault et al. 1997). This homology among members of the FCP1 family of proteins is not restricted to just the ΨΨΨDXDX(T/V)ΨΨ phosphatase motif but extends further to other conserved stretches of amino acids (see Fig. 1A). We suggest, therefore, that the FCP-homology region is the catalytic domain of a new eukaryotic protein phosphatase family. The small
molecule phosphatases and phosphotransferases share the $ΨΨΨDXDX(T/V)ΨΨ$ motif (in less highly conserved form; see Fig. 1A) but are otherwise less closely related to the members of the FCP1 family. Interestingly, the FCP1 family includes the recently described putative tumor suppressor protein HYA22 (Ishikawa et al. 1997) and the OS4 protein (Su et al. 1997) which is frequently co-amplified with CDK4 in human sarcomas. We speculate that these proteins and other members of the FCP1 family have protein phosphatase activity.

Our findings are in apparent disagreement with earlier reports which indicated that Fcp1p is not able to dephosphorylate the CTD on its own but rather needs an additional protein fraction for activity (Chambers and Kane 1996; Archambault et al. 1997). In these earlier experiments, however, the amount of Fcp1p in the reaction was very small, and the Fcp1p had been renatured after SDS-PAGE. The expression and subsequent purification of large quantities of recombinant Fcp1p allowed us to use much more Fcp1p in phosphatase assays. In the experiments described here, 40% dephosphorylation of the CTD occurred at an Fcp1p : RNAPII ratio of 0.2 and 90% dephosphorylation occurred at approximately equimolar concentrations of Fcp1p and RNAPII. In the presence of TFIIF and another stimulatory factor found in yeast extracts (Chambers and Kane 1996; Archambault et al. 1997), the turnover rate for Fcp1p is much higher and dephosphorylation of the CTD by Fcp1p is highly catalytic. Human FCP1 is also highly catalytic on mammalian RNAPII in the presence of TFIIF (Chambers et al. 1995), but, unlike yeast Fcp1p, may need no other factor besides TFIIF for high catalytic activity.
(Archambault et al. 1998b; Marshall et al. 1998). We have shown that inactivation or
destruction of Fcp1p leads to an increase in the amount of hyperphosphorylated CTD
and a decrease in the amount of hypophosphorylated CTD in vivo. This demonstrated
that Fcp1p is necessary for dephosphorylation of the CTD and is likely to be the major
CTD phosphatase. Nevertheless, it is still possible that other phosphatases in yeast also
have the CTD as a substrate.

**Fcp1p is required for transcription by RNAPII**

The synthetic lethality that we observed between *fcp1* mutations and reduced
levels of RNAPII provided evidence that Fcp1p is a positive regulator of transcription by
RNAPII. This result also implies that the effect of Fcp1p on RNAPII is probably
sufficient to explain why Fcp1p is required for yeast cell growth. Consistent with this
genetic interaction between RNAPII and Fcp1p, we have observed that Fcp1p can bind
purified RNAPII directly *in vitro* (M.S.K. and J.Greenblatt, unpublished data). Many
other components of the transcription machinery have been shown to interact genetically
with RNAPII (Hampsey 1998), but the only two gene products in which mutations had
previously been shown to be synthetically lethal with mutations in RNAPII are the
general transcription factor TFIIB (Berroteran et al. 1994) and the transcription
elongation factor TFIIS (Archambault et al. 1998a). Yeast with mutations in another
elongation factor, Spt5p, have a severe synthetic growth phenotype when combined with
mutations in Rpb1p (Hartzog et al. 1998).
There is a large decline in the total amount of poly(A)⁺ mRNA after a shift of an *fcp1-l* mutant to the non-permissive temperature. This is sufficient to explain why FCP1 is an essential gene, although we cannot exclude the possibility that Fcp1p has substrates other than RNAPII *in vivo*. Genome-wide expression analysis revealed a broad Fcp1p requirement for transcription by RNAPII *in vivo*. Although Fcp1p has not been considered a general transcription factor because it is not required for transcription *in vitro* (Orphanides et al. 1996), our data indicate that the effect of Fcp1p on transcription by RNAPII *in vivo* is comparable to that of previously characterized general transcription factors such as Kin28p, Srb4p, and Tfa1p (Holstege et al. 1998). The mRNA levels for 69% of all yeast genes in an *fcp1-l* mutant fall at about the same rate as those in an *rbp1-l* mutant. The very rapid decline of these mRNAs implies that Fcp1p is inactivated very soon after an *fcp1-l* mutant is shifted to 37°C. The majority of genes whose mRNAs decline rapidly may be the ones which are most sensitive to a decline in the pools of RNAPII holoenzyme containing unphosphorylated RNAPII. The genes whose mRNA levels are less affected or unaffected by the inactivation of Fcp1p may be the ones whose promoters compete most effectively for lowered pools of RNAPII holoenzyme. Alternatively, some of these genes may be less dependent on the phosphorylation state of the RNAPII CTD.

Little is known about CTD dephosphorylation during the transcription cycle. Since two general transcription factors, TFIIB and TFIIF, regulate CTD phosphatase activity *in vitro* (Chambers et al. 1995), and human FCP1 is present in an RNAPII holoenzyme (Archambault et al. 1998b), Fcp1p activity may be regulated at the level of
the preinitiation complex. Our studies here have shown that CTD dephosphorylation by Fcp1p is important for mRNA synthesis. It appears most likely that this Fcp1p requirement for transcription is caused by the failure of RNAPII with a hyperphosphorylated CTD to form preinitiation complexes. It is not clear, however, when the CTD becomes dephosphorylated in order for RNAPII to be able to reinitiate transcription. It is thought that the CTD must remain phosphorylated during chain elongation in order for elongation to be processive (Bentley 1995) and, perhaps, to support RNA processing, a process which is also associated with the CTD (Neugebauer and Roth 1997). Either Fcp1p cannot act on elongating RNAPIIO ternary complexes or, perhaps, other polypeptides associated with the elongating RNAPIIO, such as the recently described elongator complex (Otero et al. 1999), may inhibit CTD dephosphorylation by Fcp1p. Inhibition of CTD phosphatase activity may enhance chain elongation by RNAPII and, indeed, the HIV-1 Tat protein, which stimulates elongation by RNAPII during the transcription of HIV-1 DNA, can bind human FCP1 (J. Archambault and J. Greenblatt, unpublished data) and inhibit its CTD phosphatase activity (Marshall et al. 1998). It is unclear whether CTD phosphatase is activated and the CTD is eventually dephosphorylated before or after transcription termination occurs. Indeed, it is conceivable that dephosphorylation of the CTD by Fcp1p during transcription may trigger a loss of transcriptional processivity that leads to transcription termination.
Materials and Methods

Plasmid constructions

The *FCP1* locus on *XhoI-XbaI* fragment was excised from a Ycp50 library clone and ligated into *XhoI-XbaI* cut pRS316 (URA3⁺) (Sikorski and Hieter 1989) to generate pMK86. A *KpnI-SacI* fragment from pMK86 was inserted into pRS314 (TRP1⁺) to create pFK1. This plasmid was used for site-directed mutagenesis with the Quick Change system (Stratagene) to create pFK4 [pRS314-*fcpl-1* (R250A, P251A)], pFK7 [pRS314-*fcpl-2* (L177A, L181A, H187A)], pMK58 [pRS314-*fcpl*(D180E)], pMK59 [pRS314-*fcpl*(D180N)], pMK61 [pRS314-*fcpl*(D182E)] and pMK63 [pRS314-*fcpl*(D182N)]. All mutations were confirmed by DNA sequencing. Plasmid pMK90 [pRS314-*fcpl*-3 (Δ2-134)] was constructed by replacing a *SacI-PstI* (at amino acid 134) fragment in pFK1 with a PCR-amplified *SacI-PstI* fragment containing the *FCP1* promoter region. Plasmid pMK71 [pRS306-*LEU2* promoter- *RPBI* ORF (to 300nt past a unique *SpeI* site)] was made by inserting PCR amplified DNA from pDJ58 (David B. Jansma, unpublished data) into the *HindIII* and *NotI* sites of pRS306 (URA3⁺). A PCR-amplified *FCP1* ORF digested with *NcoI* and *XhoI* was inserted into pFastBac HTb (Gibco) to give pMK87. pMK87 was used for site directed-mutagenesis to create pMK88 [pBac-Fcp1 (D180E)] and pMK89 [pBac-Fcp1 (D182E)].

Yeast Strains
All strains are isogenic derivatives of W303-1A, with the exception of YMK88 and YMK89 which are derivatives of ZMY60 (Moqtaderi et al. 1996). YMK16 has a chromosomal deletion of the FCP1 gene (fcpIΔ::LEU2) that is complemented by the plasmid pMK86 containing genomic FCP1 on the URA3 CEN/ARS plasmid pRS316. Introducing plasmids pFK1, pFK4, pFK7, and pMK90 into YMK16 and subsequent counterselection of the URA3 marker on SC plates lacking Leu and Trp and containing 5-fluoroorotic acid (5-FOA) generated strains YMK18 (FCPI), YMK20 (fcpI-1), YMK26 (fcpI-2), and YMK28 (fcpI-3), respectively.

YMK32 (MATα fcpIΔ::LEU2 pLEU2-RPBI leu2-3, 112 or LEU2 [FCP1 URA3 CEN/ARS]) was constructed in two steps. YMK16 (MATα fcpIΔ::LEU2 [FCP1 URA3 CEN/ARS]) was mated to YF1745 (MATα rpbl::ADE2 [RPBI TRP1 CEN/ARS], David Jansma, unpublished results). A haploid progeny from this cross (MATα fcpIΔ::LEU2 rpbl::ADE2 [FCP1, URA3 CEN/ARS][RPBI, TRP1 CEN/ARS]) was then mated to YF1971 (MATα LEU2 promoter-RPBI LEU2, (Archambault et al. 1996)). YMK32 is a pink (LEU2 promoter-RPO21) 5-FOA sensitive (fcpIΔ::LEU2 [FCP1 URA3 CEN/ARS]) haploid progeny of this diploid. Plasmids pFK1, pFK4, pFK7 and pMK90 were introduced into YMK32, and subsequent counterselection of the URA3 marker on SC plates lacking Leu and Trp and containing 5-FOA generated strains YMK36, YMK38, YMK40 and YMK42, respectively.

YMK88 and YMK89 were created by digesting pMK71 with SpeI and transforming the linearized plasmid into YMK2, which contains a chromosomal ANB-
URLF-FCPl fusion gene, and the isogenic control strain YMK1, which contains pRS314. Transformants were selected on SC plates lacking leucine and containing 5-FOA. Replacement of a portion of the RPBI upstream region by the LEU2 promoter was confirmed by PCR.

Production of recombinant Fcp1p

Plasmids pMK87, pMK88 and pMK89 were transformed into DH10Bac E.coli cells (Gibco) for transposition into baculovirus. Recombinant bacmids, verified by PCR, were transfected into insect cells. Recombinant baculovirus stocks were then used to infect SF9 cells for 48 hr. To purify recombinant proteins, the infected cells were resuspended in binding buffer (20 mM HEPES pH7.6, 5 mM imidazole, 0.5 M NaCl, 10% glycerol, 5 mM β-mercaptoethanol, protease inhibitors (Complete™, Boehringer Mannheim)) and broken by sonication. Insoluble material was removed by centrifugation and the supernatant was incubated with Ni-agarose beads (Qiagen). The beads were washed twice with binding buffer and bound proteins were eluted with elution buffer (20 mM HEPES pH 7.6, 400 mM imidazole, 0.5 M NaCl, 10% glycerol, 1 mM DTT, protease inhibitors). The eluted proteins were dialyzed into phosphatase buffer (20 mM TrisCl pH 7.9, 0.1 M NaCl, 10% glycerol, 0.5 mM DTT, 0.1 mM EDTA) and loaded onto a MonoQ FPLC column. The column was washed with phosphatase buffer containing 0.3 M NaCl and eluted with a gradient from 0.3 M NaCl to 0.55 M NaCl in
phosphatase buffer. Fractions containing purified recombinant Fcp1p were identified by SDS-PAGE and staining with Coomassie blue.

Whole-Cell extracts and Immunoblot Analysis

Protein extracts from yeast strains grown in liquid culture were prepared essentially as described (Foiani et al. 1994). Briefly, 15 OD₆₀₀ units of mid-log phase cultures were centrifuged, resuspended in 200 μl of 20%TCA and disrupted by vortexing for 2 min with 200 μl acid-washed glass beads (Sigma). The supernatant was removed and the beads washed three times with 200 μl of 5%TCA. Pooled supernatants were centrifuged 10 min at 4000 rpm. The pellet was air-dried and then resuspended in 400 μl of 2x SDS gel sample buffer and 200 μl of 1 M Tris base. The extracts were boiled and insoluble material was removed by centrifugation as described above.

Various amounts of protein extract from each sample were subjected to SDS-PAGE, blotted onto nitrocellulose membranes (Schleicher and Schuell), and probed with mouse monoclonal antibody 8WG16, polyclonal rabbit antibodies against phosphorylated CTD (a gift from D. Bentley, University of Colorado), and affinity-purified rabbit polyclonal antibodies against Thioredoxin-Fcp1p (amino acids 456-667) and GST-Tfb1p (amino acids 1-136) fusion proteins. Immunoblot conditions with 8WG16 were varied so as to detect (Valay et al. 1995) or not detect (Patturajan et al. 1998a) hyperphosphorylated RNAPIIIO (the former conditions probably dephosphorylate the
RNAPII CTD on the membrane to make it detectable by 8WG16). Detection was by enhanced chemiluminescence (Pierce).

RNA Analysis

One μg of total yeast RNA purified by hot phenol extraction (Ausubel 1997) was transferred to GeneScreen (NEN Life Sciences) by using a slot blotter as recommended by the manufacturer, fixed to the membrane by UV cross-linking using a Stratalinker (Stratagene), and baked for two hours at 80°C under vacuum. Prehybridization was for 4 hours at 37°C in 5x SSPE, 10x Denhardt's, 1% SDS, 10% dextran sulfate (Ausubel 1997). The blot was hybridized overnight in the same solution at 37°C with oligo-(dT)_{30} end-labeled with $^{32}$P by T4 polynucleotide kinase. After hybridization, the blot was washed 4 times for 20 min with 2x SSPE, 0.1% SDS and then twice with 1x SSPE, 0.1% SDS at room temperature, air-dried and autoradiographed.

Measurement of run-on transcription

Yeast cells were grown in YPD at 26°C to an OD_{600} of 0.5, then transferred to 37°C or maintained at 26°C for 45 min. Triplicate aliquots of cells were centrifuged and permeabilized with sarkosyl as described (Elion and Warner 1986). Labeling was carried out in 20 μl reactions containing 10μCi [α-$^{32}$P]-UTP for 7 min at 25°C. The reactions were then diluted to 100 μl with 10 mM EDTA. 20 μl aliquots were spotted onto filter papers and TCA-precipitable radioactivity was determined.
Whole-Genome Transcription Analysis

*FCP1* and *fcp1-1* strains were grown at 30°C in YPD to an OD$_{600}$ of 0.5, diluted with an equal volume of prewarmed YPD (44°C) and then transferred to 37°C. After 45 min at 37°C total RNA was isolated, and the genomic expression profile of each strain was determined in two independent experiments per strain using Affymetrix Gene Chip arrays as described previously (Apone et al. 1998; Holstege et al. 1998). Detailed information on HDA technology and data analysis can be found on the world wide web at http://www.wi.mit.edu/young/CTD_phosphatase.html in the section titled Study Design.

**In vitro** Phosphatase Assays

The activity of Fcp1p towards p-nitrophenylphosphate (pNPP) was measured in reactions containing 10 mM KOAc, 10% glycerol, 50 mM TrisCl pH7.9, 0.1 mM EDTA, 0.5 mM DTT, 10 mg/ml pNPP, 10 mM MgCl$_2$. Reactions (50μl) were incubated for 1 hour at 30°C and stopped by the addition of 50 μl 1 M sodium carbonate. Absorbance was measured at 410nm. One enzyme unit produces 1 pmol nitrophenol per hour.

For dephosphorylation assays on the CTD, RNAPII was $^{32}$P-labeled with TFIIH as described (Chambers and Kane 1996) using TFIIH generously provided by P. Tran and M. Sayre, and then purified using a Biospin 30 column (BioRad) to remove unincorporated [$\gamma-^{32}$P]-ATP. Labeled RNAPII (150ng) was incubated with various amounts of recombinant wild-type and mutated Fcp1p proteins in 25 μl reactions for 60
min at 30°C in the phosphatase assay buffer described above. Reactions were stopped by the addition of SDS gel sample buffer and analyzed by SDS-PAGE. The gel was dried and exposed to a Phosphor Imager screen.
Chapter 3

A Motif Shared by TFIIF and TFIIB Mediates their Interaction with the RNA Polymerase II CTD Phosphatase Fcp1p in *Saccharomyces cerevisiae*

A version of this Chapter was published in “Molecular and Cellular Biology” Vol. 20, pp. 7438-7449 (2000).

I performed the experiments shown in Figures 1, 2, 3, 4, 5, 7, 8. Lisa D. Simon, who was working as a summer student with me, assisted in the experiments in Fig. 2A and Fig. 8A and James G. Omichinski did the experiment in Fig. 6.
Abstract

Transcription by RNA polymerase II (RNAPII) is accompanied by cyclic phosphorylation and dephosphorylation of the carboxy-terminal heptapeptide repeat domain (CTD) of its largest subunit. We have used deletion and point mutations in Fcp1p, a TFIIF-interacting CTD phosphatase, to show that the integrity of its BRCT domain, like its catalytic domain, is important for cell viability, mRNA synthesis, and CTD dephosphorylation in vivo. Although regions of Fcp1p carboxy-terminal to its BRCT domain and at its amino-terminus are not essential for viability, deletion of either of these regions affects the phosphorylation state of the CTD. Two portions of this carboxy-terminal region of Fcp1p bind directly to the first cyclin-like repeat in the core domain of the general transcription factor TFIIB as well as to the RAP74 subunit of TFIIF. These regulatory interactions with Fcp1p involve closely related amino acid sequence motifs in TFIIB and RAP74. Mutating the Fcp1p-binding motif KEFGK in the RAP74 (Tfg1p) subunit of TFIIF to EEFGE leads both to synthetic phenotypes in certain fcpl tfgl double mutants and a reduced ability of Fcp1p to activate transcription when it is artificially tethered to a promoter. These results suggest strongly that this KEFGK motif in RAP74 mediates its interaction with Fcp1p in vivo.
Introduction

Transcription initiation by RNA polymerase II (RNAPII) requires the assembly of a multiprotein complex at the promoter. This complex consists of RNAPII, general transcription factors and the SRB (suppressor of RNA polymerase B)/mediator proteins which are involved in the positive and negative regulation of transcription. Assembly of this preinitiation complex can be made to occur in a stepwise fashion *in vitro* (Buratowski et al. 1989) but most transcriptional initiation events in *S. cerevisiae* appear to use a preassembled RNAPII holoenzyme containing most of the essential factors (Thompson and Young 1995).

TFIIH is a general transcription factor that has an associated helicase as well as protein kinase activity (Feaver et al. 1993; Schaeffer et al. 1993; Feaver et al. 1994). One of the major targets for phosphorylation by TFIIH in the transcription initiation complex is the unique carboxy-terminal domain (CTD) of the largest subunit of RNAPII. This CTD consists of tandem repeats of the consensus sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser, repeated 52 times in human RNAPII and 26 or 27 times in *S. cerevisiae* (Allison et al. 1985; Corden et al. 1985). Roles for the CTD during transcription initiation, promoter clearance, chain elongation and transcript processing have been suggested (Neugebauer and Roth 1997; Bentley 1999). RNAPII molecules with a hypophosphorylated CTD are preferentially recruited to the initiation complex *in vitro* (Lu et al. 1991; Chesnut et al. 1992) whereas the elongating RNAPII *in vivo* usually has a hyper-phosphorylated CTD (Cadena and Dahmus 1987; O'Brien et al. 1994). Purified RNAPII holoenzyme contains
hypo-phosphorylated forms of the CTD (Koleske and Young 1994), whereas a purified form of the elongating RNAPII complex has a CTD that is heavily phosphorylated (Otero et al. 1999). Therefore it seems that the transcription cycle involves cyclical phosphorylation and dephosphorylation of the RNAPII CTD (Dahmus 1996).

Gene-specific roles have also been suggested for the CTD. Phosphorylation of the CTD by the kinase activity of the RNAPII holoenzyme component Srb10p before RNAPII binds to the promoter prevents the formation of productive transcription initiation complexes at repressed promoters (Hengartner et al. 1998). In addition, although the CTD is essential for yeast cell growth, likely reflecting its requirement for mRNA synthesis, at least several genes in *S. cerevisiae* can be transcribed by RNAPII molecules lacking the CTD (Lee and Lis 1998; McNeil et al. 1998).

Dephosphorylation of the CTD must occur in order to regenerate the non-phosphorylated form of the enzyme that appears to be recruited to promoters. A CTD phosphatase activity was originally purified from HeLa cells and subsequently from *S. cerevisiae* (Chambers and Dahmus 1994; Chambers and Kane 1996). The activity of this HeLa cell CTD phosphatase is stimulated by the RAP74 subunit of the general transcription factor TFIIF and this stimulation can be inhibited by TFIIB (Chambers et al. 1995). Partial cDNAs encoding the human protein were originally identified in a screen for RAP74-interacting proteins (Archambault et al. 1998b). The C-terminal domain of human RAP74 that interacts with the human CTD phosphatase FCP1 is necessary and sufficient for RAP74-mediated stimulation of CTD phosphatase activity *in vitro* (Archambault et al. 1998b). The homologous *FCP1* gene is essential in *S. cerevisiae* and
the yeast Fcp1 protein also interacts directly with the RAP74 subunit of yeast TFIIF (Archambault et al. 1997). The phosphatase catalytic domain of Fcp1p resembles similar domains found in a number of other database proteins of unknown function and has been designated the FCP homology domain (FCPH) (Archambault et al. 1997; Kobor et al. 1999). This domain contains the phosphatase motif \( \Psi^3 \Psi^2 \Psi^1 D \Psi^2 X^3 (T/V) \Psi^1 \Psi^2 \) (\( \Psi = \) hydrophobic residue) at its catalytic center (Archambault et al. 1997; Cho et al. 1999; Kobor et al. 1999). This motif is characteristic of a new family of small molecule phosphotransferases and phosphohydrolases (Collet et al. 1998) and is essential for Fcp1p to function in \textit{S. cerevisiae} (Kobor et al. 1999). It is different from the phosphatase motifs of the three classified protein phosphatase families (Barford et al. 1998) and, therefore, Fcp1p might be the founding member of a new class of eukaryotic protein phosphatases. Recombinant human or yeast Fcp1p can dephosphorylate the CTD (Marshall et al. 1998; Cho et al. 1999; Kobor et al. 1999) and the artificial substrate p-nitrophenylphosphate \textit{in vitro} (Kobor et al. 1999). Genome-wide expression studies show that Fcp1p is generally required for transcription by RNAPII in \textit{S. cerevisiae} (Kobor et al. 1999). The human CTD phosphatase was shown to function in recycling RNAPII \textit{in vitro} (Cho et al. 1999).

In this study, we further delineate regions of yeast Fcp1p that are functionally important \textit{in vivo} with the goal of understanding what factors influence CTD phosphatase. The integrity of the BRCT domain in Fcp1p is essential for function. We demonstrate a direct interaction between Fcp1p and TFIIB and show that the binding
sites in Fcp1p for TFIIB and RAP74 are very similar. We also identify related amino acid sequence motifs in TFIIB and RAP74 that are involved in the binding of TFIIB and RAP74 to Fcp1p. Strains with mutations in RAP74 are used to show that RAP74 utilizes this motif to interact with Fcp1p in vivo. As well, we show that Fcp1p is able to strongly activate transcription when tethered to a promoter by a heterologous DNA binding domain, suggesting that Fcp1p interacts with RNAPII holoenzyme and recruits it to the promoter. This transcriptional activation by Fcp1p is mediated in part by its interaction with RAP74, a known component of the yeast RNAPII holoenzyme (Kim et al. 1994; Koleske and Young 1994).
Results

Regions of Fcp1p required for function \textit{in vivo}

We showed previously that the catalytic FCPH domain of Fcp1p is essential for its function \textit{in vivo}. We also found that cells of a strain expressing a truncated Fcp1p that lacks the carboxy-terminal amino acid residues 627-732 are viable and that the same truncated protein is capable of functioning as a CTD phosphatase \textit{in vitro} (Archambault et al. 1997). To determine how much of the carboxy-terminus of Fcp1p was dispensable for cell viability, a series of Fcp1p’s that contained carboxy-terminal deletions was constructed with the choice of truncation site based on consideration of the predicted boundaries of the BRCT domain (amino acids 499-593) (see Fig. 1A). This domain is found in a number of proteins that are involved in cell cycle checkpoint control in response to DNA damage (Bork et al. 1997), but its importance for CTD phosphatase activity is not clear. Each truncated form of Fcp1p was then tested by using a plasmid-shuffling protocol to replace a wild-type copy of \textit{FCP1} with a mutated version of the gene (Fig. 1B and M.S.K and J. Greenblatt, unpublished data for two other mutants). A form of the \textit{FCP1} gene encoding a protein that ended just after the BRCT domain at amino acid 594 (allele \textit{fcp1-5}) supported viability, whereas a gene encoding a protein that ended at amino acid 557, which is in the middle of the two predicted BRCT sub-domains (Bork et al. 1997), did not (Fig. 1B). These results suggested that the integrity of the BRCT domain in Fcp1p is important for the protein to function \textit{in vivo}. 
Figure 1. Domains of Fcp1p required for function in vivo

A) Diagram showing predicted domain boundaries of Fcp1p and FCP1 alleles used in this figure. The ability of these constructs to support yeast cell growth during the plasmid shuffling experiments shown in Fig. 1B and Fig. 1C is summarized on the right.

B) Strains with a wild-type copy of FCP1 on the URA3 CEN/ARS plasmid pRS316 and a chromosomal FCP1 deletion were transformed with TRP1 CEN/ARS plasmids carrying the wild type FCP1 gene, no FCP1 gene or fcp1 genes encoding Fcp1p with a deletion of amino acid residues 595-732 or 558-732. The ability of these mutant alleles to complement the FCP1 chromosomal deletion was tested by plasmid shuffling on synthetic complete plates lacking Leu and Trp and containing 5-fluoroorotic acid (5-FOA) to counterselect the UR3 marker at 22°C.

C) FCP1 alleles encoding proteins that have either an amino-terminal deletion, a carboxy-terminal deletion or a simultaneous deletion of both termini were tested for their ability to complement a chromosomal deletion of FCP1 by a plasmid shuffling assay at 22°C.
We also examined the growth of strains expressing Fcp1p lacking amino acids 2-133 (fcpl-3) or lacking both termini of the protein (Fig. 1C). Interestingly, the gene expressing the form of Fcp1p that lacks both termini was not able to complement the chromosomal FCPl deletion when tested in the plasmid-shuffling assay (Fig. 1C). This was a surprising result because the fcpl-3 and fcpl-5 strains grew normally (M.S.K. and J. Greenblatt, unpublished data; see also Figs. 1C and 7A).

The BRCT domain is essential for mRNA synthesis and CTD dephosphorylation

To investigate further the importance of the BRCT domain in Fcp1p, we used site-directed mutagenesis to change its most conserved residue, namely the tryptophan at position 575, to alanine. This residue is within the second predicted conserved region of the BRCT domain (Bork et al. 1997). We originally selected this residue exclusively based on its conservation among different BRCT domains because the structural and functional knowledge of BRCT domains was very limited when we began this study. Yeast strains carrying this allele, which we named fcpl-4, were viable at 30°C but did not grow at 37°C (Fig. 2A). To determine whether the W575A mutation in strains with the fcpl-4 allele caused a defect in mRNA synthesis, as do mutations in the catalytic FCPH domain, we shifted logarithmically growing cells to the non-permissive temperature and measured the levels of poly(A)' mRNA. There was a major decrease in the amount of poly(A)' mRNA as soon as one hour after the temperature shift (Fig. 2B). Consistent with this, we also found that the W575A mutation caused a defect in the dephosphorylation of RNAPII in vivo (Fig. 2C). In this experiment, we analyzed the
Figure 2. The BRCT domain is essential for mRNA synthesis and CTD dephosphorylation

A) A strain with a W575A point mutation in the most conserved residue of the BRCT domain of Fcp1p is viable but has a temperature-sensitive phenotype. Yeast strains with wild type FCP1 or the fcp1-4 mutation on TRPI CEN/ARS plasmids in an fcp1Δ::LEU2 background were grown on synthetic complete plates lacking Leu and Trp at 30°C and 37°C.

B) Strains carrying the fcp1-4 allele have a severe defect in the synthesis of poly(A)*mRNA after shift to the non-permissive temperature. Total cellular RNA was prepared from FCP1 and fcp1-4 strains at various times after shift to 37°C. A 32P-labeled oligo-dT probe was hybridized to 1 μg of total slot-blotted RNA (Kobor et al. 1999).

C) The W575A mutation interferes with dephosphorylation of RNAPII. Whole cell extracts prepared from FCP1 and fcp1-4 strains at various times after shift from 30°C to 37°C were Western blotted with monoclonal antibody G2 to estimate the relative amounts of hyperphosphorylated and hypophosphorylated forms of the Rpb1p subunit. The same blots were also probed with antibodies against Fcp1p and Vma1p.
relative amounts of hyperphosphorylated (RNAPII o) and hypophosphorylated
(RNAPII a) forms of the largest subunit of RNAPII, Rpb1p, by probing western blots of
yeast extracts prepared at various time points after the shift to 37°C with a monoclonal
antibody directed against a region of Rpb1p outside of the CTD. We were able to clearly
distinguish the two forms of Rpb1p in these strains. Consistent with previous results
(Patturajan et al. 1998a), we observed an increase in the amount of RNAPIIO and a
decrease in the amount of RNAPIIA even in wild type strains after a shift to 37°C (Fig.
2C). However, this did not have an effect on mRNA synthesis as judged by the amount
of poly(A)⁺mRNA in these strains at the permissive temperature (Fig. 2B). At the
permissive temperature there was a larger amount of Rpb1p in the fcp1-4 mutant strain
than in the wild type strain. Interestingly, a much higher proportion of the Rpb1p in the
fcp1-4 strain was present in either the hyperphosphorylated form or in a variety of
intermediate phosphorylated forms even before transfer to the non-permissive
temperature; at this time the amount of the mutant Fcp1p in the cell was normal (Fig.
2C). A detailed analysis of the abnormal CTD phosphorylation pattern in the mutant
strain at the permissive temperature will be described elsewhere (M.S.K. and J.
Greenblatt, unpublished data; see Chapter 5). These results strongly suggested that the
BRCT domain is important for the CTD phosphatase activity of Fcp1p in vivo.
Moreover, the hypophosphorylated RNAPII A form completely disappeared in the fcp1-
4 strain upon shift to 37°C; at this temperature the mutant Fcp1p was partly degraded.
The amount of the vacuolar H⁺-ATPase Vma1p, which we used as a loading control, did
not change significantly (Fig. 2C). There was a very good correlation between the decrease
in poly(A)$^+$ mRNA and the change in the phosphorylation pattern of Rpb1p, giving further support to our earlier finding that inactivation of CTD phosphatase leads to a rapid shutdown of mRNA synthesis by RNAPII.

**Fcp1p interacts directly with the general transcription factor TFIIB**

Previous studies in the human system have shown that TFIIB can regulate CTD phosphatase activity *in vitro* (Chambers et al. 1995). Consistent with this, a two-hybrid screen using human FCP1a (Archambault et al. 1998b) as a bait led to the identification of human TFIIB as an interacting protein (J. Langlois and J. Greenblatt, unpublished data). Therefore, we examined whether the corresponding yeast proteins were able to interact with each other *in vitro*. Fcp1p could bind to recombinant full-length yeast TFIIB in an affinity chromatography experiment (Fig. 3A) and Fcp1p also bound a portion of the TFIIB that is present in yeast whole cell extract (M.S.K. and J. Greenblatt, unpublished data). TFIIB consists of an amino-terminal zinc-ribbon domain (amino acids 1-100) and a core domain (amino acids 100-345) containing two cyclin-related repeats (Hisatake et al. 1993) (Fig. 3B). When a series of GST fusion proteins containing various domains of TFIIB was used as ligands in affinity chromatography experiments, only the fusion proteins containing either the complete core domain or the first cyclin-like repeat were able to bind Fcp1p (Fig. 3C). This suggested that Fcp1p interacts with the first cyclin-related repeat of TFIIB.
Figure 3. Fcp1p interacts with the first cyclin-like repeat in TFIIB

A) Fcp1p binds to the recombinant full length TFIIB. Purified recombinant His_{6}-tagged yeast TFIIB was bound to Ni-Agarose (Qiagen) microcolumns at the indicated concentrations and \(^{35}\)S-labeled Fcp1p made by \textit{in vitro} transcription and translation was tested for binding. After washing, the bound protein was eluted with high salt buffer containing 1 M NaCl and visualized by autoradiography after SDS-PAGE.

B) Diagram showing predicted domain boundaries of yeast TFIIB

C) The first cyclin-like repeat of the TFIIB core domain mediates the binding of Fcp1p. The indicated GST-yTFIIB fusion proteins were purified from bacteria and coupled to Glutathione-Sepharose columns at a concentration of 4 mg/ml, then \(^{35}\)S-labeled Fcp1p was tested for binding. The eluted Fcp1p protein was analyzed by SDS-PAGE and visualized by autoradiography.
Similar regions of Fcplp bind TFIIB and RAP74

We used the TFIIB core domain as an affinity chromatography ligand to test which regions of yeast Fcplp are involved in the Fcplp-TFIIB interaction. Versions of Fcplp with carboxy-terminal and amino-terminal deletions were tested for their ability to bind the immobilized GST-TFIIB (100-345). We had previously shown that two adjacent regions of yeast Fcplp containing amino acid residues 457-666 and 667-732 can each bind yeast RAP74 independently (Archambault et al. 1997) (see also Fig. 4B). These two regions were also able to bind to GST-TFIIB (100-345) (Fig. 4A). In this case, a thioredoxin (TRX) fusion protein containing amino acid residues 457-666 of Fcplp bound GST-TFIIB (100-345) more strongly than a TRX-Fcpl (667-732) fusion protein. We were unable to detect binding when we tried to further sub-divide these regions of Fcplp, suggesting that more extensive deletions influence the proper folding of these regions (M.S.K and J.Greenblatt, unpublished data). Of the proteins with carboxy-terminal deletions that we tested, the deleted protein lacking amino acid residues 667-732 of Fcplp was only slightly compromised in binding TFIIB, whereas removal of amino acids 627-732 completely abolished the interaction (Fig. 4A). We next used these same constructs to perform a more detailed mapping of the binding regions for RAP74 in Fcplp in order to compare them to the binding regions for TFIIB and also to the regions dispensable for viability. Similar to what was observed with TFIIB, Fcplp (1-626) had no detectable binding to RAP74 (Fig. 4B). The TRX-Fcpl (457-666) fusion protein bound somewhat less strongly to RAP74 than the TRX-Fcpl (667-732) fusion protein or the full length protein. It appeared that Fcplp (1-666) bound much less strongly to
Figure 4. TFIIB and RAP74 bind in a similar way to Fcp1p

Various $^{35}$S-labeled portions of Fcp1p or thioredoxin (TRX)-Fcp1 fusion proteins made by transcription and translation in vitro were chromatographed over affinity columns containing the indicated concentrations of either GST-yTFIIB (100-345) (Fig. 4A) or GST-yRAP74 (649-735) (Fig. 4B). GST was used as a control in all binding experiments. The columns were washed and bound proteins were eluted with reduced glutathione, analyzed by SDS-PAGE and visualized by autoradiography.

Fig. 4C summarizes the relative binding strengths of various portions of Fcp1p to GST-yTFIIB (100-345) and GST-yRAP74 (649-735) presented in Fig. 4A and Fig. 4B.
RAP74 than to TFIIB when compared to the full-length protein. Therefore, we concluded that two adjacent regions in Fcplp can interact with both TFIIB and RAP74 and that the interaction with both factors is abolished when amino acid residues distal to amino acid 627 of Fcplp are deleted. One of the binding sites in Fcplp, namely amino acids 457-666, binds TFIIB about four times as strongly as it binds RAP74, whereas the other binding site in Fcplp, namely amino acids 667-732, binds RAP74 at least ten times as strongly as it binds TFIIB.

**An amino acid sequence motif common to TFIIB and RAP74 mediates their binding to Fcplp**

The similar binding patterns of yeast TFIIB and RAP74 to Fcplp prompted us to search for similar amino acid sequences in these two general transcription factors. The sequence KEFGK is present in both yeast TFIIB and RAP74 (Fig. 5A). The amino acid sequences in these regions are also similar in the TFIIB and RAP74 proteins found in *S. cerevisiae*, human, *Xenopus laevis*, *Drosophila melanogaster* and *Caenorhabditis elegans*.

In the structure of human TFIIB, this amino acid sequence (amino acids 186-196) maps to part of helix E1 at the end of the first cyclin-like repeat (Bagby et al. 1995; Nikolov et al. 1995). Previous studies have suggested that this region within yeast TFIIB is important for the formation of the TFIIB-TBP-DNA complex, and it has also been shown to participate in binding to acidic activators *in vitro* (Bangur et al. 1997; Chou and Struhl...
1997). To test whether this same region in yeast TFIIB (amino acids 198-208) was involved in the binding of Fcp1p, we constructed a K201E point mutation in Figure 5. A conserved amino acid sequence motif is involved in the binding of TFIIB and RAP74 to Fcp1p

A) Comparison of sequences from the first cyclin-related repeat of yeast TFIIB and the carboxy-terminal domain of yeast RAP74. The alignment can be extended to TFIIB and RAP74 from other species (not shown). Mutations that were created are indicated below (RAP74) or above (TFIIB) the alignment.

B) A K201E amino acid change in TFIIB reduces the binding to Fcp1p. Wild type and a K201E derivative of GST-TFIIB (100-345) were bound to Glutathione-Sepharose at the indicated concentrations and affinity chromatography with 35S-labeled Fcp1p was performed as described in the Fig. 3 legend.

C) Amino acid residues within the KEFGK motif of RAP74 are important for binding to Fcp1p. A series of amino acid changes was introduced in GST-RAP74 (649-735) as indicated in Fig. 5A, and the resulting proteins were tested for binding to 35S-labeled Fcp1p in affinity chromatography experiments. All the mutations that were tested reduced the binding of Fcp1p to GST-RAP74 (649-735).
the context of a GST-TFIIB (100-345) fusion protein. This mutation was previously shown to cause a temperature-sensitive phenotype in yeast, whereas TFIIB with a K201E/K205E double mutation was unable to support yeast growth (Bangur et al. 1997; Chou and Struhl 1997). The K201E derivative of GST-TFIIB (100-345) bound Fcp1p much less strongly than the wild type protein (Fig. 5B).

We next made a series of mutations within this same motif in the carboxy-terminal region of yeast RAP74 (amino acids 649-735). Four mutant GST-RAP74 (649-735) fusion proteins were expressed, purified and tested for the ability to bind Fcp1p in affinity chromatography experiments. Mutating one or both of the conserved lysine residues at positions 695 and 699 to glutamate had a strong effect on the ability of the resulting GST-RAP74 fusion proteins to bind Fcp1p (Fig. 5C). This effect was stronger for the double mutation (yRAP74-2) than the single mutation (yRAP74-1). We also changed both of these basic residues to the neutral amino acid alanine (yRAP74-3) and tested for the binding of Fcp1p. Again, the mutated protein was not as efficient as the wild type protein in binding Fcp1p. In order to determine whether the binding was purely dependent on the basic charge of this region, we also tested an E696K mutation (yRAP74-4) which increased the basic charge of this region. This altered protein also bound Fcp1p more weakly than wild type RAP74. These data suggested that both the charge and the sequence of this region are important for binding to Fcp1p.

It was possible, however, that the mutations we created in the Fcp1p-binding motif of RAP74 disturbed the proper folding of the C-terminal region of RAP74 that we were using in our binding assays. Therefore, we prepared an 15N-labeled sample of the carboxy-
Figure 6. RAP74 (649-735) and RAP74 (649-735) K695E/K699E are properly folded protein domains

The double mutation K695E/K699E does not affect the overall folding of the RAP74 carboxy-terminal region as shown by 2D $^1$H-$^{15}$N HSQC NMR spectra of the wild type (A) and K695E/K699E (B) proteins.
terminal domain of the RAP74 K695E/K699E protein and compared its 2D $^1$H, $^{15}$N-
HSQC NMR spectrum with that of the same portion of the wild-type protein (Fig. 6).

Excellent spectra were obtained for both the mutant and the wild-type proteins, and
almost all of the 80 expected $^1$H-$^{15}$N amide peaks could be counted. The substantial
dispersion of the chemical shifts suggested that the wild type domain had a stable three-
dimensional fold. This result therefore indicated for the first time that the carboxy-
terminus of RAP74 can form an independently folded domain. Importantly, the spectrum
of the double mutant looked very similar with only a few exceptions that probably came
from the altered amino acids or their immediate vicinity. This indicated that the overall
folding of the protein was not affected by the amino acid changes. Further evidence for
proper folding of this domain and the integrity of the mutant protein was obtained by
circular dichroism experiments (M.S.K., A.R. Davidson and J. Greenblatt, unpublished
data). Therefore, the two important lysine residues of the Fcp1p-binding motif might
interact directly with acidic residues in the carboxy-terminal domain of Fcp1p to define
the molecular interface.

Effect of mutations in the Fcp1p-binding motif of RAP74 \textit{in vivo}

The gene encoding the yeast homologue of the RAP74 subunit of human TFIIF is
called \textit{TFG1} (Henry et al. 1994). To test the importance for yeast cell growth of the
Fcp1p-binding motif of RAP74, a \textit{tfgl-2/TFG1} diploid yeast strain was constructed. The
\textit{tfgl-2} allele encodes RAP74 (K695E/K699E), which fails to bind Fcp1p \textit{in vitro} (Fig.
5C). After sporulation we obtained a viable haploid \textit{tfgl-2} strain. Growth of this strain
Figure 7. Effects in vivo of mutations in the Fcp1p binding domain of RAP74

A) The K695E/K699E mutation in the carboxy terminal domain of RAP74 causes sensitivity to the DNA damaging drug Methanosulfonic Methyl Ester (MMS). TFG1 wild type and tfgl-2 mutant strains were plated at 2000 cells per plate on synthetic complete medium plates lacking His and containing 0.01% MMS and grown at 30°C for 3 days. Control plates did not contain any MMS.

B) Synthetic phenotypes of fcp1 tfgl-2 double mutant yeast strains. Tenfold serial dilutions of TFG1 or tfgl-2 strains that also had a chromosomal deletion of FCP1 and carried the indicated FCP1 alleles on TRP1 CEN/ARS plasmids were grown for 3 days at various temperatures on SC plates lacking His, Trp and Leu. All the strains with mutant fcp1 alleles are more temperature-sensitive in the tfgl-2 background than in the TFG1 background, except for the strain carrying the fcp1-5 allele which encodes a protein that lacks the RAP74-binding site. The diagram shows the positions of the mutations in Fcp1p.

C) Western blot analysis of fcp1 tfg1 double mutant strains. Whole cell extracts prepared from strains carrying the FCP1, fcp1-1 and fcp1-4 alleles in the TFG1 and tfgl-2 background were grown at 22°C, subjected to SDS-PAGE and Western blotted with monoclonal antibody G2 to estimate the relative amounts of the hyperphosphorylated and hypophosphorylated forms of the Rpb1p subunit. A similar experiment was also performed with strains carrying the FCP1, fcp1-3 and fcp1-5 alleles that were grown at 30°C. The blots were underexposed to visualize the difference in the phosphorylation
status of the CTD of Rpb1p in the double mutant strains. The same blots were also probed with antibodies against Vma1p.
was normal in all conditions that we tested, with the exception that growth was somewhat impaired in the presence of the DNA-damaging agent methanosulfonic methyl ester (MMS) (Fig. 7A). This phenotype was also observed in fcp1 mutant strains (M.S.K. and J. Greenblatt, unpublished data), therefore providing a phenotypic link between Fcp1p and RAP74.

We also tested the effect of various fcp1 alleles on the growth of the tfgl-2 strain. Plasmids carrying the FCP1, fcp1-3, fcp1-4 and fcp1-5 alleles described above, as well as plasmids carrying previously described fcp1-1 and fcp1-2 alleles with viable point mutations in the FCP homology domain (Kobor et al. 1999), were introduced into cells of a fcp1Δ tfgl-2 strain harboring a plasmid containing URA3 and FCP1. Plasmid shuffling experiments showed that none of the fcp1 mutant alleles affected the viability of the tfgl-2 strain at 22°C. We then tested these double mutant strains for growth at various temperatures (Fig. 7B). The results showed that strains with the fcp1-1, fcp1-2 and fcp1-4 alleles containing mutations in the catalytic and BRCT domains, as well as the amino-terminal deletion allele fcp1-3, were more temperature-sensitive in the tfgl-2 background than in the TFG1 background. Importantly, the C-terminal deletion allele fcp1-5, which creates a truncated protein lacking amino acids 595-732 and unable to bind RAP74 in vitro (M.S.K and J. Greenblatt, unpublished data), was not affected by the tfgl-2 mutation that prevents RAP74 from binding to Fcp1p. These observations are consistent with our binding data, supported the hypothesis that RAP74 interacts with the carboxy-terminal portion of Fcp1p in vivo, and suggested that this interaction becomes more
important when the function of Fcp1p is weakened by particular physiological conditions or by mutations in Fcp1p.

These conclusions were further supported by western blot analysis of Rpb1p in extracts prepared from single mutant strains and fcp1 tfg1 double mutant strains grown at the permissive temperature. In most cases, there was a larger amount of Rpb1p as well as a higher proportion of intermediately phosphorylated and hyperphosphorylated Rpb1p in the double mutant strains than in strains with mutations only in Fcp1p (Fig.7C). The CTD phosphorylation status of Rpb1p was essentially identical in the single and double mutant strains only in the case of strains with the fcp1-5 allele that encodes a Fcp1p which lacks the RAP74-binding site.

**RAP74 can be a target for transcriptional activation in vivo**

A number of components of the RNAPII holoenzyme complex have been shown to be able to activate transcription when artificially tethered to a promoter via a fusion to a heterologous DNA-binding domain (Chatterjee and Struhl 1995; Farrell et al. 1996). This activation is thought to be due to enhanced recruitment of the RNAPII holoenzyme (Ptashne and Gann 1997). Since human FCP1 is a component of a human RNAPII holoenzyme complex and since Fcp1p interacts with TFIIB and RAP74, which are components of yeast RNAPII holoenzyme (Koleske and Young 1994), we tested whether yeast Fcp1p, when fused to the DNA-binding domain of LexA, could activate the transcription of a *lacZ* reporter gene containing 8 upstream LexA-binding sites *in vivo*. As shown in Fig. 8A, a LexA-Fcp1p fusion strongly activated transcription. Removing an
increasing number of amino acid residues from the carboxy-terminus of the LexA-Fcp1 fusion protein resulted in a sharp drop in transcriptional activation. LexA-yFcp1 (1-666) had only about 5% of the activation potential of the full length fusion protein, and the shorter fusion protein LexA-yFcp1 (1-626), which does not bind RAP74 or TFIIIB (Fig. 4), failed to activate transcription above the levels of the LexA DNA-binding domain alone. All these fusion proteins were expressed to comparable levels in the yeast cells (M.S.K., L. Simon and J. Greenblatt, unpublished data). Therefore, these results revealed a good correlation between the ability of Fcp1p to activate transcription in vivo when brought into the vicinity of a promoter and its ability to bind to RAP74 and TFIIIB in vitro.

We also tested smaller portions of Fcp1p for ability to activate the reporter gene as LexA-fusions (Fig. 8A). The LexA-yFcp1 (457-732) fusion protein containing both TFIIF/TFIIB interaction sites strongly activated the reporter gene, whereas LexA-Fcp1p (457-666) and LexA-Fcp1p (627-732), each of which likely contains only a simple RAP74/TFIIB interaction site, activated transcription less strongly. Interestingly, the LexA-Fcp1p (457-732) and LexA-Fcp1p (457-666) constructs contained the BRCT domain previously shown to mediate transcriptional activation by BRCA1 (Chapman and Verma 1996). The smallest portion of Fcp1p that was able to activate transcription when fused to LexA contained amino acid residues 627-732. Smaller portions of Fcp1p, although able to bind to both RAP74 and TFIIIB in vitro (see Fig. 4), were not stably expressed in yeast cells as judged by western blot analysis with anti-LexA antibodies (M.S.K., L. Simon and J. Greenblatt, unpublished data).
Figure 8. Transcriptional activation by LexA-Fcp1 fusion proteins

A) A full length LexA-Fcp1 fusion protein and a series of carboxy-terminal and amino-terminal deletion constructs were tested for their ability to activate the lacZ reporter construct pSH18-34, which has 8 LexA-binding sites upstream of the promoter. The ability to activate transcription is correlated with ability to bind RAP74 and TFIIB.

B) The interaction between Fcp1p and RAP74 is important for transcriptional activation by LexA-Fcp1p. LexA or a LexA-Fcp1 fusion protein was expressed in a TFG1 or tfg1-2 strain and transcriptional activation of reporter constructs having one, two or eight LexA-binding sites was measured. Transcriptional activation by LexA-Fcp1p is synergistic and is dependent on the interaction with RAP74. This dependence is more pronounced at higher levels of activation. For reasons of simplicity, the numbers for LexAp alone were omitted but were always below 2 Miller units. All measurements were done in triplicate and the average is given. Standard deviations were below 20% for all points.
Evidence for the importance of the Fcp1p-RAP74 interaction in the process of transcriptional activation by LexA-Fcp1p was obtained by performing the activation assays in tfgl-2 mutant cells. In this experiment, we examined the effect of the number of LexA-binding sites on transcriptional activation by LexA-Fcp1 (1-732) fusion proteins in TFG1 and tfgl-2 strains. As shown in Fig. 8B, LexA-Fcp1p could significantly activate transcription from a reporter construct with only one binding site. Transcriptional activation increased synergistically with the number of LexA-binding sites, similar to the situation with classical activators. Importantly, transcriptional activation by LexA-Fcp1p was reduced in tfgl-2 strains. This reduction became more pronounced with an increase in the number LexA-binding sites that led to a higher level of activation by the LexA-Fcp1 fusion protein. Western blot analysis showed that the LexA-Fcp1p fusion proteins were expressed to similar levels in the TFG1 and tfgl-2 strains (M.S.K. and J.Greenblatt, unpublished data). The difference in activation of the same reporter genes with either one or eight LexA binding sites by a LexA-Gal4 fusion protein was less then 10% in TFG1 versus tfgl-2 strains (M.S.K. and J.Greenblatt, unpublished data). These results indicated both that the KEFGK motif in RAP74 is a major target for transcription activation by LexA-Fcp1p and that there are other targets for Fcp1p (e.g. TFIIB) in the transcription apparatus. These results also provided further evidence that Fcp1p and RAP74 interact in vivo.
Discussion

Fcp1 is an RNAPII CTD phosphatase required for most or all mRNA synthesis in yeast. A characteristic feature of the Fcp1p is the presence of a BRCT domain, which is commonly found in proteins that are involved in checkpoint control in response to DNA damage. We demonstrated that the integrity of this domain is essential for Fcp1p to function in S. cerevisiae. Even at the permissive temperature, a strain with a point mutation in the BRCT domain of Fcp1p accumulates an excessive proportion of hyperphosphorylated RNAPII. Upon shift to the non-permissive temperature the hypophosphorylated form of RNAPII disappears and mRNA production is shut down. This further supports our earlier finding, obtained by using yeast strains carrying the \textit{fcp1-1} and \textit{fcp1-2} alleles, that Fcp1p is responsible for dephosphorylating the CTD \textit{in vivo} and that failure to do so can result in a shutdown of RNAPII transcription (36). A recent X-ray structure of the BRCT domain of the human DNA repair protein XRCC1 showed that its two predicted subdomains form one compact domain consisting of a four-stranded parallel beta-sheet surrounded by three alpha helixes with extensive intramolecular contacts (Zhang et al. 1998). Based on this structure, it is very likely that our partial deletion of the BRCT domain of Fcp1p, which was not able to complement a chromosomal \textit{FCPI} deletion, profoundly affected the overall fold and structure of the domain. The conserved tryptophan residue that we altered to create a temperature-sensitive mutant forms part of the highly conserved hydrophobic pocket and makes contacts with a number of other important residues. Our mutation might create a version
of the BRCT domain that is partly unfolded and quickly unfolds upon shift to the non-permissive temperature, leading to degradation of the mutant protein. BRCT domains have been shown to be dimerization domains involved in the formation of either homodimers (Soulier and Lowndes 1999) or heterodimers (Taylor et al. 1998) and can also interact with proteins like RNA helicase A that do not contain BRCT domains (Anderson et al. 1998). Interestingly, mutating the conserved Trp residue in XRCC1 to Asp did not affect the interaction with DNA ligase III (Taylor et al. 1998). It will be very important to identify the interacting protein partner(s) of this essential domain of Fcp1p. No other protein within the general transcription machinery is known to have a BRCT domain.

In addition to the known interaction between Fcp1p and RAP74 (Archambault et al. 1997), our studies revealed a direct interaction with the general transcription factor TFIIB, which had been shown previously to inhibit RAP74-stimulated CTD phosphatase activity (Chambers et al. 1995). We discovered that helix E1 within the first cyclin-like repeat in the core domain of yeast TFIIB mediates the interaction with Fcp1p. This region was implicated previously in the formation of the TFIIB-TBP-DNA complex and binding to the acidic activation domain of VP16. Mutations within this region cause a temperature-sensitive phenotype and have an effect on basal transcription in vitro (Bangur et al. 1997; Chou and Struhl 1997). However these mutated TFIIB proteins retained the ability to respond to the acidic activator Gal4-VP16 in vitro and in vivo, suggesting that this region is important for basal but not activated transcription (Bangur et al. 1997; Chou and Struhl 1997). The additional interaction of this region with CTD
phosphatase is intriguing, and more work will be necessary to distinguish between effects
causd by the inability to form the TFIIB-TBP-DNA complex or failure to interact with
Fcp1p.

Although we did not further address the function of TFIIB in the regulation of
CTD phosphatase activity, we showed that there are two independent binding sites for
TFIIB within the carboxy-terminal region of Fcp1p. These binding sites are very similar
to the binding sites for RAP74, although the stronger binding site for RAP74 lies within
amino acids 667-732 whereas the stronger binding site for TFIIB lies within amino acids
457-666. A truncated version of Fcp1p that ends at amino acid 626 and lacks binding
sites for TFIIB and RAP74 is still active in a CTD phosphatase assay in vitro
(Archambault et al. 1997).

Nevertheless, our in vivo deletion analysis demonstrated that a protein that fails
to interact with both TFIIB and RAP74 is able to support growth. CTD phosphatase
interacts directly with RNAPII ((Chambers et al. 1995); M.S.K. and J.Greenblatt,
unpublished data) and there might be additional interactions with components of the
RNAPII holoenzyme which might suffice in vivo to bring CTD phosphatase into the
vicinity of its substrate and compensate for the lack of strong RAP74 - Fcp1p and TFIIB
- Fcp1p interactions in most circumstances. Another explanation for this finding is the
 possibility that an unidentified CTD phosphatase distinct from Fcp1p can compensate
for the loss of Fcp1p interaction with TFIIB and RAP74. TFIIB, SRB10 and CTDK-1
are distinct CTD kinases which might phosphorylate the CTD on different residues
during the transcription cycle. Therefore the existence of distinct CTD phosphatases specific for different phosphorylated residues cannot be excluded.

We identified a previously unrecognized short amino acid sequence of high similarity between TFIIB and RAP74 that maps to a carboxy-terminal domain of RAP74 and helix E1 in at the carboxy-terminal end of the first cyclin repeat of TFIIB (Bagby et al. 1995; Nikolov et al. 1995). The degree of amino acid conservation within this motif is particularly high among the TFIIB and RAP74 proteins of *S. cerevisiae*, human and *Xenopus laevis*, although less so for *Drosophila melanogaster* and *Caenorhabditis elegans*, suggesting that it serves an important biological function. We showed that this motif mediates the interactions of TFIIB and RAP74 with Fcp1p. Our data suggest that the charge distribution of this region is a major component of the interactions and are consistent with earlier findings that RAP74 can stimulate CTD phosphatase activity and that this stimulation can be inhibited by TFIIB (Chambers et al. 1995). Our data imply that TFIIB may compete with RAP74 for binding to Fcp1p. However, we have not been able so far to determine whether the interactions between RAP74 and TFIIB with the carboxy-terminal half of Fcp1p are mutually exclusive or whether the three proteins can form a ternary complex.

Just as deletion of the portion of Fcp1p that binds TFIIB and RAP74 in strains that carry the *fcp1-5* allele has little effect on cell growth, so does the *tfg1-2* mutation in RAP74 that prevents it from binding to Fcp1p. However, the sensitivity of the *tfg1-2* mutant to the DNA damaging agent MMS provides an indirect phenotypic link to *FCP1*. The failure of CTD phosphatase to interact with RAP74 may lead to loss of viability
when there is severe DNA damage or other stresses which we have not yet discovered.

Importantly, double mutant analysis strongly supports a functional interaction between RAP74 and the carboxy-terminal region of Fcp1p in vivo. Strains carrying a mutant fcp1 allele are more temperature sensitive in the tfgl-2 background than in the TFG1 background, except in the case of fcp1-5, which lacks the RAP74-binding sites. These results are consistent with the hypothesis that stimulation of CTD phosphatase activity by RAP74 occurs in vivo and becomes more important when CTD phosphatase activity is weakened by mutations in Fcp1p outside its RAP74-binding region. These results also imply that the KEFGK motif in RAP74 mediates its interaction with Fcp1p in vivo.

Consistent with this, we found a higher proportion of intermediately phosphorylated and hyperphosphorylated Rpb1p in double mutant strains except in the case where the carboxy-terminal RAP74-binding region of Fcp1p is deleted.

Our studies also revealed that the amino-terminal region of Fcp1p has some function in vivo. Firstly, deletion of both the amino-terminal and carboxy-terminal regions of Fcp1p is lethal. Although we do not know whether the resulting protein is unstable or non-functional, an fcp1-3 tfgl-2 strain carrying an amino-terminal deletion of Fcp1p is temperature-sensitive. Taken together, these data suggest that the amino-terminal region of Fcp1p has an important function that is only revealed when Fcp1p cannot interact with RAP74. Strains that lack either the carboxy-terminus or the amino-terminus of Fcp1p have a higher proportion of intermediately phosphorylated Rpb1p.

Our observation that transcriptional activation by LexA-Fcp1p is strongly reduced in the tfgl-2 mutant strain provided further evidence that RAP74 interacts with
Fcp1p in vivo. The artificial recruitment of RNAPII holoenzyme by a LexA-Fcp1p fusion protein and subsequent activation of the reporter construct transcription may occur because Fcp1p can associate with RNAPII holoenzyme complexes (Archambault et al. 1998). We have observed that a LexA-Fcp1p fusion is able to complement a chromosomal fcp1 deletion, suggesting that it can perform the normal cellular functions of Fcp1p (M.S.K., L.Simon. and J.Greenblatt, unpublished data). The substantial effect of the tfg1-2 mutation on activation by LexA-Fcp1p indicated that this activation probably involves a direct interaction with RAP74. The fact that portions of Fcp1p that do not contain the phosphatase catalytic domain are able to activate transcription indicates that activation by LexA-Fcp1p is independent of CTD phosphatase activity and suggests that recruitment of RNAPII holoenzyme is the most likely mechanism for this activation process. The acidic carboxy-terminal region of Fcp1p could also act as a bona fide transcriptional activator. Interestingly, previous studies have shown an important role for human RAP74 in transcriptional activation by serum response factor (SRF) (Joliot et al. 1995), and an interaction between the activation domain of the model activator Gal4-VP16 and RAP74 has been reported (Zhu et al. 1994). Although the interaction sites for SRF, VP16 and Fcp1p on RAP74 do not overlap, it is tempting to speculate that some endogenous yeast activators might act at least partially through interaction with RAP74 in vivo. The fact that the effect of the tfg1-2 mutation on activation by LexA-Fcp1p decreases when fewer LexA-binding sites are present suggested that LexA-Fcp1p also interacts with some protein other than RAP74. This protein could be TFIIB or some other component of the RNAPII holoenzyme.
CTD phosphatase activity may be regulated partly at the level of transcriptional initiation, as is suggested by its physical and functional interactions with TFIIIF and TFIIIB. It has been reported that CTD phosphatase activity is necessary to recycle RNAPII after a round of transcription is completed in vitro (Cho et al. 1999), and it is possible that both TFIIIF and TFIIIB play a role in the temporal regulation of this process. Alternatively, it is also possible that CTD phosphatase is active at a different step during the transcription cycle. Although TFIIIF binds directly to RNAPII and helps to assemble RNAPII into the preinitiation complex, it can also interact with elongating RNAPII and stimulate its rate of chain elongation (Flores et al. 1989; Price et al. 1989; Izban and Luse 1992). A role for TFIIIB in steps after holoenzyme recruitment to the promoter has also been proposed (Ranish et al. 1999). One or more factors in HeLa nuclear extract affect the ability of Fcp1p to dephosphorylate the CTD in transcription elongation complexes in vitro (Cho et al. 1999; Lehman and Dahmus 2000), and the HIV Tat protein, which stimulates chain elongation by RNAPII binds to human Fcp1p and inhibits its CTD phosphatase activity (Marshall et al. 1998). In light of our studies reported here, it may be important that the interaction between HIV-Tat protein and human FCP1 involves the BRCT domain of human FCP1 (J.Archambault and J.Greenblatt, unpublished data). It is conceivable, therefore, that CTD phosphatase activity is differentially regulated at more than one stage of the transcription cycle. As well, it cannot be excluded that Fcp1p has relevant substrates other than the CTD in vivo. In this regard, it is interesting that a number of the general transcription factors, including RAP74, have been reported to be phosphoproteins (Sopta et al. 1985; Ohkuma and Roeder 1994), and phosphorylation of
RAP74 has been suggested to be important for HIV-Tat mediated stimulation of transcriptional elongation (Zhou et al. 2000b). In addition, it is possible that Fcp1 has functions that are not related to its phosphatase activity, as suggested by the finding that human FCP1 can act as an elongation factor in vitro independent of its catalytic function (Cho et al. 1999).
Materials and Methods

Plasmids

Plasmids for the *in vivo* analysis of *FCPI* deletions were constructed with a two-step strategy using the pRS series of *CENARS* plasmids (Sikorski and Hieter 1989). pFK1 [pRS314-*FCPI*(1-732)] was digested with *Pst*I, which cuts at codon 134 of *FCPI*, and *Xho*I, which cuts in the pRS314 polylinker downstream of the *FCPI* transcription termination sequence. PCR products obtained by using primer pairs MK44 (5'-GGG CTG CAG ATG CCT TCG ATG TG TAC C-3') and MK46 (5'-GGG CTC GAG CTA GAT TAA CGT GTA GGG TTT TTC ATC C-3') or MK44 and MK47 (5'-GGG CTC GAG CTA TTT ATA AGT GCT AGG GTT CTT GG-3') (stop codons are underlined) were cut with *Pst*I and *Xho*I and inserted into the pRS314-*FCPI* vector backbone to create plasmids containing versions of *fcpl* coding for amino acids 1-594 and 1-557, respectively, but lacking the transcription termination sequence. The termination sequence was subsequently inserted into these constructs cut with *Xho*I and *Kpn*I by using a PCR product obtained with the primer pair MK48 (5'-GGG CTC GAG CYC AGA TGC GT ATC TTT CC-3') and MK49 (5'-GGG GGT ACC AGT ACT TGT TGA GTA TTT AGG GG-3') and digested with *Kpn*I and *Xho*I to give pMK96 [pRS314-*fcpl*-5 (1-594)] and pMK97 [pRS314-*fcpl*-6 (1-557)]. Plasmid pMK98 [(pRS314-*fcpl*-7 (134-594)] was constructed by excising the *Pst*I / *Kpn*I fragment from plasmid pMK96 and inserting it into pMK90 [pRS314-*fcpl*-3 (134-732)] cut with *Pst*I and *Kpn*I.
pLS2 [pRS314 - fcp1-4 (W575A)] was constructed by site directed mutagenesis with the quick change mutagenesis kit (Stratagene) using primers MK87 (5'-GTT CAC CCA GAT GCG ATA TTC GAA TGT TTG G-3') and MK88 (5'-CCA AAC ATT CGA ATA TCG CAT CTG GGT GAA C-3') (the codon for the changed amino acid is underlined) and pFK1 [pRS314-FCPI] as the template. The mutation was confirmed by DNA sequencing.

Plasmids encoding amino acid residues 1-120, 100-345, 100-240 and 210-345 of yeast TFIIB fused to Glutathione-S-transferase (GST) were constructed by inserting PCR fragments cut with BglII and XhoI into the BamHI and XhoI sites of pGEX-4T-1 (Pharmacia). In each case, plasmid pET19d-His yTFIIB (gift from A. Emili) was used as the template and the amplifying primers were as follows: MK5 (5'-GGG AGA TCT ATG ATG ACT AGG GAG AGC-3') and MK11 (5'-GGG CTC GAG CTA ATC CAT CAC ATT TTT TCC TTG-3') to create pMK91 [pGEX-4T-1-yTFIIB (1-120)]; MK7 (5'-GGG AGA TCT ACC ACG GAT ATG AGA TTC AC-3') and MK9 (5'-GGG CTC GAG CTA TTT CTT TCC AAC GCC CGG-3') to create pMK92 [pGEX-4T-1-yTFIIB (100-345)]; MK7 and MK10 (5'-GGG CTC GAG CTA GGG TAT ATA AGT TAG GTT TTG-3') to create pMK93 [pGEX-4T-1-yTFIIB (100-240)]; MK8 (5'-GGG AGA TCT ATG AAG AAC ATT TTA AGA GGC-3') and MK9 to create pMK94 [pGEX-4T-1-yTFIIB (210-345)]. Stop codons are underlined. The expression vector for mutant GST-TFIIB K201E (pMK95) was constructed by PCR using primers MK7 and MK9 with plasmid pQE/yIIIB K201E (a gift from A. Ponticelli) as the template. The
PCR product was cut with BglII and XhoI and inserted into pGEX-4T-1 cut with BamHI and XhoI.

Expression vectors for mutant GST-γRAP74 proteins were derived from pJA728 [pGEX-3X-γRAP74 (649-735)] using the quick change site directed mutagenesis kit (Stratagene). pMK65 [pGEX-3X-γRAP74 (649-735) K695E] was constructed by using primers MK55 (5′-GGC AAA GTC AAT ATC GAA GAA TTC GGA AAG TTC ATC-3′) and MK56 (5′-GAT GAA CTT TTC GAA TTC TTC GAT ATT GAC TTT GCC-3′); pMK66 [pGEX-3X-γRAP74 (649-735) K695E, K699E] was constructed by using primers MK61 (5′-GGC AAA GTC AAT ATC GAA GAG GAA TTC ATC AGA AAG-3′) and MK62 (5′-CTT CTG ATG AAT TCT CCA AAC TCT TCG ATA TTG ACT TTG CC-3′); pMK67 [pGEX-3X-γRAP74 (649-735) K695A, K699A] was constructed by using primers MK70 (5′-GTC AAA GAA TTC GGA AAG TTC ATC AGA AAG-3′) and MK71 (5′-CTT CTG ATG AAT TCT CCA AAC TCG GCC TGC TTC ATC AGA AG-3′) and MK72 (5′-GTC AAT ATC AAA AAG TTT GGA AAG TTC-3′) and pMK68 [pGEX-3X-γRAP74 (649-735) E696K] was constructed by using primers MK72 (5′-GTC AAT ATC AAA AAG TTT GGA AAG TTC-3′) and MK73 (5′-GAA CTT TCC AAA CCT TTT GAT ATT GAC-3′). Codons encoding the mutated amino acids are underlined.

Plasmids for in vitro transcription/translation were constructed as follows:

pMK96 [pRS314-fcp1-5 (1-594)] and pMK101 [pRS314-fcp1-8 (1-626)] were digested with SpeI (which cuts at codon 457 of FCP1) and XhoI (cutting after the stop codon) and the insert was ligated into plasmid JA782 cut with SpeI and XhoI to give plasmids pMK102 [pET23d-HA-Fcp1 (1-594)] and pMK103 [pET23d-HA-Fcp1 (1-626)],
respectively. The other plasmids used for in vitro transcription / translation have been described previously.

Plasmids for the chromosomal integration of TFGI alleles were constructed in a two step strategy. In the first step, the TFGI termination sequence was amplified from genomic DNA by PCR with primers MK131 (5'-GGG GGG GGA TCC GTT AGT TTA TAA TGT TAT GTA C-3') and MK132 (GGG GGG CTC GAG CTG GAA GAG AAT ACT TAA GAG-3'), digested with BamHI and XhoI, and inserted into the BamHI and XhoI sites of the integrating vector pRS303, which has a HIS3 marker for selection in yeast. The carboxy-terminus of TFGI was amplified from plasmid JA728, in the case of wild type TFGI, or plasmid pMK66, in the case of the tfgl-2 mutant, by PCR with primers MK133 (5'- GGG GGG TCT AGA GGA ATC CAC AGA CGA CAA AAG CTG TAG ATA GTA GTA ATA ATG CAT CGA ATA CAG TGC CTT CGC C-3') and MK134 (5' - GGG GGG GGA TCC CTA CCC GGG AGC GTA GTC TGG AAC GTC GTA TGG GTA CTC TTT CTT TTA TTA CAT GTG GTC ATT GCC-3') which also encodes a HA tag (in italics). The PCR products were digested with BamHI and XbaI and inserted into the pRS303-TFGI terminator plasmid cut with XbaI and BamHI to give pMK99 [pRS303-TFGI (636-735) HA] and pMK100 [pRS303-tfgl-2 (636-735) HA K695E/K699E].

Expression vectors for LexA-Fcp1 fusion proteins were derivatives of pEG202 (Ausubel 1997). Various portions of FCP1 were amplified from pMK86 [pRS316-FCP1] using PCR. pJA815 [LexA-yFcp1 (1-732)] was constructed by using primers LexA-start (5'-CCC CAG ATC TCC ATG GCT TAC CCA TAC GAT G-3') and yFIP1-stop (5'-
CCC AGA TCT GAT ACG GCA TCT GAG CTG AGC TGC TAA TC-3'). pMK72 [LexA-yFcp (1-666)] was constructed by using primers MK5 (5'-CCC AGA TCT CCA TGA CCA CAC AAA TAA GGT C-3') and MK4 (5'- CCC CTC GAG CTA GTC GTG GTC GTC ATC TTC-3'); pMK73 [LexA-yFcp1(1-644)] was constructed by using primers MK5 and MK69 (5'-GGG CTC GAG CTA TAA CCA TGA AGT ACC AGC AGC-3'); pMK74 [LexA-yFcp1 (1-626)] was constructed by using primers MK5 and MK68 (5'-GGG CTC GAG CTA ATG CTG TTG TTC CTG GGT C-3'); pMK75 [LexA-yFcp1 (457-732)] was constructed by using primer MK1 (5'-CCC AGA TCT CCG TTG ATG ACG ATG ATG AAC-3') and MK3 (5'-CCC CTC GAG CTA ATC ATC CAG CAT ATC C-3'); and pMK76 [LexA-Fcpl (457-666)] was constructed by using primers MK1 and MK4. Stop codons are underlined. The PCR products were cut with BglII and XhoI and inserted into pEG202 cut with BamHI and XhoI. pMK77 [LexA-yFcp1 (626-732)] was constructed by using primers MK101 (5'-CCC CGG ATC CCC TTG ACA TCA CAA GAA AAT CTA AAT TTA TTC-3') and MK3; pMK78 [LexA-yFcp1 (645-732)] was constructed by using primers MK102 (5'-CCC CGG ATC CCC AAC AAT GAC GAC GAT GAA GAT ATT CC-3') and MK3; pMK79 [LexA-yFcp1 (667-732)] was constructed by using primers MK103 (5'-CCC CGG ATC CCC GAC GAA AGT GAT GAC GAA AAC AAC TCG-3') and MK3. The PCR products were cut with BamHI and XhoI and inserted into the BamHI and XhoI sites of pEG202.

Protein purification
GST-fusion proteins were expressed in *E. coli* DH5α cells. Cells were grown at 30°C to an OD$_{600nm}$ of 0.4 and induced with 1 mM IPTG. After 3 h induction, the cells were harvested and resuspended in 1 M buffer A (10 mM TrisHCl pH7.9, 1 mM EDTA, 1 mM DTT, 1 M NaCl). The cells were lysed by sonication and the supernatant after centrifugation was mixed with Glutathione-Sepharose 6B (Pharmacia). The beads were washed four times with 1 M buffer A and then two times with 0.1 M buffer A. The bound proteins were eluted with 0.1 M ACB (10 mM HEPES pH7.9, 1 mM EDTA, 1 mM DTT, 10% glycerol, 100 mM NaCl) containing 50 mM reduced glutathione and dialyzed into 0.1 M ACB. The purified proteins were analyzed by SDS-PAGE and bound to fresh Glutathione-Sepharose 6B at the indicated concentrations for affinity chromatography experiments.

The carboxy-terminal-domains of wild type RAP74 and the K695E/K699E double mutant were overexpressed as GST fusions in *E. coli* BL21 grown in $^{15}$N-labeled M9 minimal medium. The fusion protein was first purified using Glutathione-Sepharose. The fragments were then cleaved from the GST with Factor X and purified using ion exchange chromatography on an SP-Sepharose Fast Flow column. After this initial purification step, the samples were concentrated to 1 mM and then dialyzed into the NMR buffer (10 mM Sodium Phosphate pH 6.0, 0.25 mM EDTA, 1 mM DTT).

**Protein-protein interaction assays**
GST and various GST-TFIIB and GST-RAP74 fusion proteins were coupled to Glutathione-Sepharose 6B at the indicated concentrations. The columns (40 μl) were equilibrated with 200 μl of 0.1 M ACB containing 5 mg BSA/ml and next with 400 μl 0.1 M ACB containing 1 mg BSA/ml. Columns were loaded with 20 μl rabbit reticulocyte lysate from the TNT transcription/translation system (Promega) programmed with 0.4 μg of the various plasmid DNAs, which had been diluted tenfold in 0.1 M ACB containing 1 mg BSA/ml. The columns were washed with 400 μl of 0.1 M ACB and the bound proteins were eluted in 120 μl 0.1 M ACB containing 50 mM glutathione. Thirty μl aliquots of the eluted proteins were analyzed by SDS-PAGE and autoradiography.

Western Blot Analysis

Protein extracts from yeast strains grown in synthetic complete medium lacking the appropriate amino acids were prepared by glass bead lysis in the presence of trichloroacetic acid as described previously (Kobor et al. 1999). After SDS-PAGE and transfer to a nitrocellulose membrane, protein analysis was performed using either the monoclonal antibody G2 (a generous gift from V. Svetlov and R. Burgess) directed against a conserved domain within the largest subunit of yeast RNAPII, an affinity purified polyclonal antibody against Fcp1p (Kobor et al. 1999), or a monoclonal antibody against the vacuolar H⁺-ATPase VMA1p (Molecular Probes) using standard procedures.

NMR studies
All NMR experiments were performed at 25°C on a Varian Inova 600MHz spectrometer equipped with a pulse field gradient unit and triple resonance ($^1$H/$^13$C/$^{15}$N) probe with an actively shielded z gradient. Sensitivity enhanced gradient 2d ($^1$H, $^{15}$N) HSQC spectra were recorded with a $^{15}$N sweep width of 1600 Hz centered at 118 ppm, and 64 complex points t1. An 8000 Hz 1H sweep width centered at 4.753 was recorded with 512 complex points t2.

Yeast Strain Construction

All strains are derivatives of W303-1A (ade2-1 can1-100 trpl-1 leu2-3,112 his3-11,15 ura3-1 ssd1-d2). Integration of the TFG1 wild type and tfg1-2 mutant alleles was done by cutting pMK99 and pMK100, respectively, with EcoRI and transforming W303-1A diploid cells. Yeast cells were transformed with linearized plasmids using the LiAc procedure and integrants were selected on synthetic complete (SC)-His plates. Integration was confirmed by PCR and haploid spores were obtained using a hydrophobic sporulation protocol (Ausubel 1997), creating strains YMK202 and 204. The mutation was confirmed by sequencing PCR-amplified genomic DNA.

YMK202 and YMK204 were mated with YMK16 (Kobor et al. 1999). Diploids were selected on SC-His/Ura/Leu plates. Haploid strains YMK215 (TFG1) and YMK217 (tfg1-2) were obtained after sporulation and hydrophobic spore enrichment. These strains also contain a chromosomal knockout of the FCP1 gene and are kept alive by pMK86 [pRS316-FCP1].
YMK215 and YMK217 were used for plasmid shuffling experiments with plasmids pFK1 [pRS314-FCPI], pFK4 [pRS314-fcp1-1], pFK7 [pRS314-fcp1-2], pMK90 [pRS314-fcp1-3], pLS2 [pRS314-fcp1-4] and pMK96 [pRS314-fcp1-5]. The transformed cells were grown at 22°C for 4 days on SC plates lacking Trp, Leu and His and then streaked on SC plates containing 5-fluoroorotic acid at 22°C for 4 days to counterselect the URA3 marker.

YMK211 and YMK212 are TFG1 and tfg1-2 strains in which the HIS3 marker originally used for the integration at the TFG1 locus was disrupted by the TRP1 marker. This was done by transforming the linearized plasmid pH6 (Cross 1997) into YMK202 and 204, respectively. Transformants were selected on SC-Trp plates, and only transformants that did grow on SC plates lacking Trp but not on SC plates lacking His (indicating that the HIS3 gene at the TFG1 locus was disrupted by the TRP1 gene and not the original his3-11,15 allele in W303) were selected. This manipulation was necessary for measuring the transcriptional activation by LexA-Fcp1p fusion proteins in TFG1 versus tfg1-2 strains because the reporter plasmids that we used have a HIS3 marker.

β-galactosidase assays

Liquid culture assays to measure β-galactosidase produced by yeast strains that were transformed with the LexA operator-lacZ fusion plasmid pSH18-34 and various LexA-Fcp1p derivatives were performed by using cells permeabilized with sarkosyl essentially as described (Kippert 1995). Cells were grown to mid-logarithmic phase in SC
medium lacking Ura and His. The number of cells used for the various LexA-yFcp1 fusion proteins was adjusted in order to obtain reliable \( \text{OD}_{420\text{nm}} \) readings. For each measurement, \( \beta \)-galactosidase activity was determined from three independent cultures and average values are given.

For Fig. 8B, strains YMK211 (\textit{TFG1}) and YMK212 (\textit{tfgl-2}) were transformed with plasmid JA816 [LexA-Fcp1 (1-732)] or JA821 [LexA-Stop] as well as one of the \textit{lacZ} reporter plasmids pRB1840 (1 LexA-binding site, Origene), pJK103 (2 LexA-binding sites, Origene), or pSH18-34 (8 LexA-binding sites), and enzyme activity was measured as above except that cells were grown in SC medium lacking Ura, His and Trp.
Chapter 4

Opposing Roles of the RNA Polymerase II CTD Phosphatase Fcp1p and the CTD Kinase Srb10p in *Saccharomyces cerevisiae*

I performed the experiments shown in Figures 1, 2, 3A and 4. Lisa D. Simon did the experiment in Fig. 3B as a summer student under my supervision. Paul Jorgensen made the yeast strain that I used in plasmid shuffling experiments to obtain the strains used in Fig. 1C.
Abstract

The cyclin-dependent kinases Kin28p and Srb10p are involved in phosphorylating the carboxy-terminal domain (CTD) of RNA polymerase II (RNAPII). We found that strains with mutations in the CTD phosphatase Fcp1p have much more severe growth defects when Kin28p is mutated. In contrast, temperature-sensitive fcp1 mutants survive at higher temperatures and are less sensitive to cell cycle-checkpoint inducing drugs when Srb10p is inactivated by the srb10D290A mutation. Inactivating Srb10p also partly reverses the over-accumulation of the hyperphosphorylated Rpb1p subunit of RNAPII that occurs in fcp1 mutant cells at the permissive temperature. These results suggest that Fcp1p and Srb10p have competing roles in regulating the phosphorylation state of the RNAPII CTD in vivo.
Introduction

Accurate and regulated initiation of transcription requires the assembly of RNA polymerase II (RNAPII), general transcription factors (GTF’s) and the SRB (suppressor of RNA polymerase B) or mediator complex at the promoter (Orphanides et al. 1996). Although assembly of this complex can be made to occur in a stepwise fashion in vitro (Buratowski et al. 1989), a preassembled RNAPII holoenzyme containing RNAPII, mediator and some GTFs may be used for most transcriptional initiation events in the yeast Saccharomyces cerevisiae (Thompson and Young 1995).

The largest subunit of RNAPII contains a unique carboxy-terminal domain (CTD) consisting of tandem repeats of the consensus sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser, repeated 52 times in human RNAPII and 26 or 27 times in yeast (Allison et al. 1985; Corden et al. 1985). Purified yeast RNAPII holoenzyme contains the hypophosphorylated form of the CTD (Koleske and Young 1994) and, consistent with this, hypophosphorylated RNAPII is preferentially recruited into preinitiation complexes in vitro (Lu et al. 1991; Chesnut et al. 1992). Elongating transcription complexes in S. cerevisiae contain RNAPII molecules that are hyperphosphorylated on the CTD and tightly associated with the elongator complex (Svejstrup et al. 1997; Otero et al. 1999). The RNAPII CTD appears to have important roles during initiation, promoter clearance, chain elongation and mRNA processing, and these disparate roles may be coordinated in part by phosphorylation of the CTD (Neugebauer and Roth 1997; Hirose and Manley 2000; Proudfoot 2000).

At least three protein complexes participate in the phosphorylation of the CTD in
S. cerevisiae. TFIIH is a 9-subunit general transcription factor with associated CTD kinase and DNA helicase activities (Feaver et al. 1991; Feaver et al. 1993; Schaeffer et al. 1993). The Kin28p/Ccllp kinase/cyclin pair in TFIIH is responsible for phosphorylation of the CTD (Feaver et al. 1994). This TFIIH-associated CTD kinase is particularly active within the preinitiation complex and is required for mRNA synthesis by RNAPII (Holstege et al. 1998). TFIIH seems to phosphorylate primarily Ser5 of the CTD heptapeptide repeat YSPTSPS during transcription initiation and promoter clearance (Komarnitsky et al. 2000). Phosphorylation of the CTD by the Kin28p kinase is involved in the subsequent recruitment to the transcription complex of mRNA capping enzymes (Komarnitsky et al. 2000; Rodriguez et al. 2000; Schroeder et al. 2000).

In contrast, the Srb10p/Srb1lp kinase/cyclin pair, which can associate with RNAPII holoenzyme, is capable of phosphorylating the CTD before formation of the preinitiation complex (Hengartner et al. 1998), and this phosphorylation might inhibit the formation of the preinitiation complex on promoters that are regulated by Srb10p. Genome-wide expression studies have revealed a negative effect of the Srb10p kinase activity on the expression of a limited set of genes (Holstege et al. 1998). Srb10p-mediated phosphorylation of the transcriptional activator Gal4p also plays a role in transcriptional activation by Gal4p (Hirst et al. 1999).

The CTDK-1 CTD kinase consists of CTK1, CTK2 and CTK3 gene products that are not essential in S. cerevisiae, although deletions of these genes lead to slow growth and cold-sensitive phenotypes (Lee and Greenleaf 1991; Sterner et al. 1995). Yeast strains in which the CTK1 gene is deleted have CTDs that are less phosphorylated, although
some phosphorylation still remains. Whereas TFIIH and Srb10p/Srb11p are components of some RNAPII holoenzyme complexes purified from yeast and human cells, CTDK1 is not found in these complexes (Sterner et al. 1995).

The CTD phosphatase FCP1 has been purified from human and yeast cells (Chambers and Dahmus 1994; Chambers and Kane 1996; Archambault et al. 1997; Archambault et al. 1998b; Cho et al. 1999). The HeLa cell CTD phosphatase is stimulated by the RAP74 subunit of TFIIF and this stimulation can be inhibited by TFIIB (Chambers et al. 1995). FCP1 is an essential gene in S. cerevisiae, and the yeast Fcp1p protein interacts directly with RAP74 and TFIIB through a common KEFGK motif present in these factors (Archambault et al. 1997; Kobor et al. 2000). Fcp1p has a FCP homology domain (FCPH) that is involved in its catalytic activity and contains a ΨΨΨDXDX(T/V)ΨΨ (Ψ = hydrophobic residue) motif that is essential for both CTD phosphatase activity in vitro and for cell viability in vivo (Kobor et al. 1999). Fcp1p also contains a functionally important BRCT domain, which is found in some proteins involved in cell cycle checkpoint control in response to DNA damage (Bork et al. 1997; Kobor et al. 2000). Fcp1p is necessary for dephosphorylation of the CTD in vivo, and yeast cells with temperature-sensitive fcp1 mutations in the FCPH or BRCT domain have profound defects in mRNA synthesis at the non-permissive temperature (Kobor et al. 1999; Kobor et al. 2000).

In this study, we investigated the relationship between Fcp1p and the two RNAPII holoenzyme-associated CTD kinases Kin28p and Srb10p. We found that
mutating Kin28p or Srb10p in cells carrying mutant fcp1 alleles had opposite effects on yeast cell growth. Double mutant fcp1 kin28 strains have much more severe growth defects than fcp1 or kin28 single mutants. In contrast, inactivating Srb10p partly suppresses the phenotypes of fcp1 mutants. Strains with abnormal phosphorylation of the CTD caused by mutating FCP1 are very sensitive to the cell cycle checkpoint-inducing drugs hydroxyurea and methanosulfonic methyl ester, and this sensitivity is partly suppressed by inactivation of Srb10p. Conversely, the invasive growth phenotype caused by inactivation of the Srb10p kinase is substantially reversed in fcp1 srb10D290A double mutant strains. These and other observations show that Fcp1p and Srb10p have opposing effects on the overall phosphorylation state of the RNAPII CTD.
Results

Genetic Interaction between \textit{KIN28} and \textit{FCP1}

\textit{KIN28} and \textit{FCP1} are essential genes, and so we used a plasmid shuffling protocol to examine the phenotypes of strains with conditional mutations in both genes. A strain was constructed that had chromosomal deletions of \textit{KIN28} and \textit{FCP1} as well as the \textit{kin28-16} temperature sensitive allele (Cismowski et al. 1995) on a \textit{TRP1 CEN/ARS} plasmid and wild-type \textit{FCP1} on a \textit{URA3 CEN/ARS} plasmid. We then introduced plasmids carrying wild-type \textit{FCP1} or the \textit{fcpl-1}, \textit{fcpl-2} and \textit{fcpl-4} temperature sensitive mutant alleles on \textit{ADE2 CEN/ARS} plasmids (Fig. 1A). When the \textit{FCP1 URA3 CEN/ARS} plasmid was removed by counter-selection with 5-fluoroorotic acid at 22°C, every \textit{fcpl} mutant allele was able to complement the chromosomal \textit{FCP1} deletion in the \textit{KIN28} and the \textit{kin28-16} backgrounds (M.S.K and J.Greenblatt, unpublished data). We then compared the effects of the various \textit{fcpl} alleles on the growth of the \textit{KIN28} and \textit{kin28-16} strains at 30°C (Fig. 1B). The results showed that strains with \textit{fcpl} mutations in the FCPH and BRCT domains were far more temperature-sensitive in the \textit{kin28-16} than in the \textit{KIN28} background. Therefore, strains with both mutations in the gene encoding the major CTD kinase, Kin28p, and the gene encoding the only known CTD phosphatase, Fcp1p, exhibited much more severe growth defects than the defects caused by mutating either gene alone.
Figure 1. Mutations in the CTD kinases Kin28p and Srb10p have opposite effects on cell growth in yeast strains with mutations in the CTD phosphatase Fcp1p.

A) Diagram outlining predicted domain boundaries of Fcp1p and positions of the mutations in Fcp1p with the corresponding allele names used in this study.

B) Synthetic effect of the kin28-16 allele on cell growth in fcp1 mutant strains. Strains with point mutations in conserved amino acid residues of Fcp1p had a stronger temperature-sensitive phenotype in the kin28-16 than in the wild-type KIN28 background. Yeast strains with FCP1 or the fcp1-1, fcp1-2, and fcp1-4 alleles on ADE2 CEN/ARS plasmids in an fcp1::LEU2 background were compared when either KIN28 was present at its chromosomal locus or the kin28-16 mutant allele was present on a TRPI CEN/ARS plasmid while the chromosomal copy of kin28 was replaced with HIS3. Cells were streaked on synthetic complete (SC) plates lacking Leu, His, Trp, and Ade and grown at 30°C.

C) Partial suppression of the temperature-sensitive phenotype of fcp1 strains by inactivation of the Srb10p kinase activity. Strains with the FCP1 or the fcp1-1, fcp1-2, and fcp1-4 alleles on TRPI CEN/ARS plasmids in a fcp1::LEU2 background and either wild type SRB10 or the srb10D290A allele in the chromosome were plated in 5-fold serial dilutions on SC plates lacking Trp and Leu and incubated at the indicated temperatures for 3 days.
A

1 173 357 499 583 732

Fcp1p

fcp1-2 fcp1-1 fcp1-4

B

FCP1

fcp1-1 fcp1-2 fcp1-4

KIN28 kin28-16

C

30°C 33°C 35°C 37°C

SRB10

wt 1-1 1-2 1-4

srb10 D290A

FCP1

wt 1-1 1-2 1-4
Mutations in the Srb10p CTD kinase can partly rescue the temperature-sensitivity caused by mutations in Fcp1p

In striking contrast to the negative genetic interaction between KIN28 and FCP1, the srb10D290A allele (Kuchin and Carlson 1998) encoding Srb10p with a catalytically inactive kinase was able to partly suppress the temperature-sensitivity caused by mutations in FCP1 (Fig. 1C). The effects of the various fcp1 mutations in the FCPH and BRCT domains on growth at various temperatures were tested in the context of either wild type SRB10 or the srb10D290A allele in the chromosome. In each case, the growth defect caused by the fcp1 mutation was less severe in the srb10D290A background than in the SRB10 background. Similar results were obtained when we used a complete deletion of the SRB10 gene instead of the srb10D290A allele (M.S.K and J.Greenblatt., unpublished data). Taken together, these data indicated that the CTD kinases Kin28p and Srb10p have opposite effects on cell growth in strains with mutations in the CTD phosphatase Fcp1p. There was little genetic interaction between CTK1 and FCP1 and a detailed analysis of functional interactions between the CTDK1 CTD kinase and Fcp1p will be published elsewhere (E.J. Cho, M.S.K., J. Greenblatt and S. Buratowski, unpublished data)

Global effect of Srb10p on CTD phosphorylation in fcp1 mutant strains.

To understand further suppression of the temperature-sensitivity of fcp1 mutants caused by inactivating Srb10p, we examined the amounts of the hyperphosphorylated (RNAPIIo) and hypophosphorylated (RNAPIIa) forms of the largest subunit Rpb1p of RNAPII. Hyperphosphorylated Rpb1p migrates more slowly in an SDS polyacrylamide
gel than hypophosphorylated Rpb1p. Extracts prepared from cells grown at the permissive temperature were western blotted with a monoclonal antibody, G2, directed against a region of Rpb1p outside the CTD (a kind gift from V. Svjetlov and R. Burgess).

The \textit{srb10D290A} mutation had little effect on the CTD phosphorylation pattern in \textit{FCP1} strains (Fig. 2, lanes 1 and 5). Consistent with our previous report (Kobor et al. 2000), we found increased amounts of total Rpb1p as well as larger amounts of hyperphosphorylated and intermediately phosphorylated Rpb1p in \textit{fcp1} mutant strains at 30°C (Fig. 2, lanes 1-4). This indicated that certain amino acid changes in the FCPH and BRCT domains of Fcp1p cause reduced CTD phosphatase activity even at the permissive temperature and reduced phosphatase activity led to elevated levels of Rpb1p.

When extracts from strains with \textit{fcp1} mutations and the additional \textit{srb10D290A} mutation were examined, we found that the amounts and CTD phosphorylation patterns of Rpb1p became more similar to the patterns in wild type cells (Fig. 2, lanes 5-8). The amounts of the third largest subunit, Rpb3p, of RNAPII did not change significantly in \textit{fcp1} or \textit{srb10D290A} strains, nor did the amounts of Fcp1p and the vacuolar H$^+$-ATPase Vma1p, which served as loading controls (Fig. 2). Similar results were obtained when we tested strains with a complete deletion of the \textit{SRB10} gene (M.S.K. and J. Greenblatt, unpublished data). Therefore, these results suggested Srb10p has an important role in phosphorylating the RNAPII CTD, which only becomes evident when CTD phosphatase activity is reduced. However, \textit{fcp1 srb10D290A} double mutant strains still contain significant amounts of hyperphosphorylated Rpb1p (Fig. 2, lanes 5-8), suggesting that another, unidentified CTD kinase also phosphorylates RNAPII.
Figure 2. The increased amount of Rpb1p and abnormal CTD phosphorylation pattern present in fcpI mutant strains is partly reversed by inactivation of Srb10p.

Whole cell extracts were prepared from cells grown at 30°C carrying the FCP1, fcp1-1, fcp1-2 and fcp1-4 alleles in the SRB10 or srb10D290A background described in Fig. 1C. The extracts were subjected to SDS-PAGE and Western blotting with monoclonal antibody G2 to estimate the relative amounts of the hypophosphorylated (IIa) and hyperphosphorylated (IIo) forms of the Rpb1p subunit. The same extracts were also tested by Western blotting for the amounts of the third largest subunit of RNAPII (Rpb3p), Fcp1p and the vacuolar H⁺-ATPase Vma1p.
Strains with mutations in Fcp1p are sensitive to cell cycle checkpoint-inducing drugs

To identify physiological effects caused by the abnormal CTD phosphorylation pattern associated with altered versions of Fcp1p, we tested yeast strains with mutations in the FCPH or BRCT domain of Fcp1p for conditional phenotypes at the permissive temperature under a variety of different conditions. No dramatic growth phenotypes were observed on various carbon sources, or upon exposing cells to various salts, or added chemicals known to promote growth defects. Only exposure to the cell cycle checkpoint-inducing drugs hydroxyurea (HU) and methanosulfonic methyl ester (MMS) caused growth defects in yeast strains carrying the fcp1-1, fcp1-2 and fcp1-4 alleles (Fig. 3A).

This specificity was interesting because the other proteins that contain BRCT domains have roles in cell cycle checkpoint control in response to DNA damage (Bork et al. 1997). The sensitivity to HU and MMS could be completely reversed by introducing a wild type copy of the FCP1 gene on a CEN/ARS plasmid into fcp1 mutant cells (M.S.K. and J. Greenblatt, unpublished data). This showed that these phenotypes were indeed caused by altered versions of Fcp1p and also that the fcp1 mutations were recessive.

Importantly, fcp1 srb10D290A double mutant strains were more resistant to both drugs (Fig. 3A). This phenotypic suppression was consistent with the ability of the srb10D290A mutation to partly reverse the effects of fcp1 mutations on the state of phosphorylation of the CTD and the amount of Rpb1p.

The genes encoding ribonucleotide reductase (RNR) are thought to play a major role in the cellular response to HU and MMS (Elledge et al. 1993). We therefore used a
Figure 3. Sensitivity of *fcp1* mutant strains to cell cycle checkpoint-inducing drugs and effects on *RNR3* gene expression.

A) The growth defects of *fcp1* strains in response to hydroxyurea and methanosulfonic methyl ester can be partly reversed by inactivation of Srb10p. Strains with mutations in *FCPI* and either wild type *SRB10* or the *srb10D290A* mutation were plated on synthetic complete plates containing 100 mM hydroxyurea (HU) or 0.075% Methanosulfonic Methyl Ester (MMS) and incubated at 30°C for 5 days. Control plates did not contain any HU or MMS and were incubated at 30°C for 3 days.

B) Induction of an *RNR3-lacZ* reporter gene in response to 100 mM HU. β-galactosidase activity in the indicated strains was measured before and 4 hours after addition of 100 mM HU. The results were averaged for at least 3 independent cultures and the standard deviation was below 20% for all points.
RNR3–lacZ reporter gene construct on a CEN/ARS plasmid (Zhou and Elledge 1992) to test the effects of HU on RNR3 expression in wild-type and mutant cells. Interestingly, we found that fcp1-l cells were significantly compromised in their ability to derepress the RNR3 promoter after exposure to HU (Fig. 3B). We also observed constitutive derepression of RNR3 in cells of the srb10D290A strain, and this resembled the previously described CRT (constitutive RNR3 transcription) phenotype (Zhou and Elledge 1992). In the fcp1-l srb10D290A strain, however, RNR3 was induced to a level comparable to that in wild-type cells (Fig. 3B). This suggested that failure to fully induce the RNR3 gene in fcp1-l mutants contributed to the HU-sensitivity of this strain.

Additionally, our results indicate that induction of the RNR3 gene might be particularly sensitive to disturbances in the normal CTD phosphorylation pattern. Consistent with this prediction, we found that yeast cells carrying RNAPII molecules with a shortened CTD also had defects in RNR3 gene expression when exposed to HU and were sensitive to this drug (M.S.K, L.D. Simon and J. Greenblatt, unpublished data).

Invasive growth caused by mutations in Srb10p can be suppressed by mutations in Fcp1p

To investigate further the genetic interaction between SRB10 and FCP1, we tested whether mutant alleles of FCP1 could suppress a phenotype caused by the srb10D290A mutation. This mutation causes haploid S. cerevisiae to grow pseudohyphae, presumably because of the derepression of certain Srb10p-regulated genes that play a role in this process (Holstege et al. 1998). Therefore, we tested the ability of haploid cells to invade
agar after 5 days of growth on YPD plates, a phenotype that is related to pseudohyphal growth in diploid cells. Invasiveness was tested by applying a gentle stream of tap water to the plate and washing off cells that had not grown into the agar. As expected, the *srb10D290A* strain was able to invade agar, unlike wild-type or *fcp1* strains (Fig. 4). In contrast, *fcp1 srb10D290A* double mutant strains were significantly less invasive (Fig. 4). This provided further evidence for a close connection between *FCP1* and *SRB10* in the regulation of gene expression.
Figure 4. Reduced Fcp1p activity can partly suppress the effects of the *srb10D290A* mutation on invasive growth

The ability of the *srb10D290A* mutant strain to invade agar is greatly reduced by mutations in *FCP1*. Cells carrying the *FCP1, fcp1-1, fcp1-2* and *fcp1-4* alleles and either *SRB10* or the *srb10D290A* allele were plated on YPD and incubated at 30°C for 5 days. Invasiveness was tested by applying tap water to wash off the non-invading cells.
Discussion

Fcp1p is the founding member of a new family of protein phosphatases and is required for dephosphorylation of the RNAPII CTD (Kobor et al. 1999). Fcp1p interacts with TFIIIB and the RAP74 subunit of TFIIIF and the interaction with RAP74 is important for proper functioning of Fcp1p in vivo (Archambault et al. 1997; Kobor et al. 2000). Fcp1p might be involved in transcriptional elongation, termination and recycling of RNAPII (Cho et al. 1999; Kobor et al. 1999; Costa and Arndt 2000; Lehman and Dahmus 2000). This study focuses on the relationship between Fcp1p and the cyclin-dependent CTD kinases Kin28p and Srb10p.

Yeast strains with mutations in fcp1 accumulate the hyperphosphorylated form of Rpb1p at 30°C, suggesting that altering conserved amino acids in Fcp1p leads to a reduction in CTD phosphatase activity. These fcp1 strains also contain a larger amount of hypophosphorylated Rpb1p than wild-type strains. In other experiments, we have found that reduced Fcp1p activity triggers a compensatory mechanism leading to increased synthesis of Rpb1p, most of which is hypophosphorylated and not assembled into functional RNAPII (see Chapter 5). At 37°C, the overproduced Rpb1p that is not assembled into RNAPII is degraded (see Chapter 5).

Several lines of evidence point to opposite functions of Fcp1p and Srb10p in transcription and CTD phosphorylation. Yeast strains with mutations in fcp1 are less sensitive to increased temperatures and to the cell cycle checkpoint-inducing drugs HU and MMS in an srb10D290A background than in an SRB10 background. This suppression
may occur because the srb10D290A mutation partly reverses the over-accumulation of hyperphosphorylated Rpb1p. Mutations in fcpl cause altered expression of more than 1000 genes even at the permissive temperature of 30°C (F.C.P. Holstege and R. Young, personal communication), and more detailed future studies could identify genes that are differentially regulated when srb10 is also mutated.

Srb10p has a major effect on phosphorylation of the CTD in purified RNAPII holoenzymes in vitro (Liao et al. 1995) even though relatively few genes are negatively controlled by Srb10p in vivo (Holstege et al. 1998). Fcp1p is thought to dephosphorylate RNAPII before or after transcriptional termination in order to ensure that most of the free RNAPII is hypophosphorylated and able to initiate transcription (Cho et al. 1999; Kobor et al. 1999). We propose that CTD dephosphorylation by Fcp1p directly opposes CTD phosphorylation by Srb10p, although this global phosphorylation of RNAPII by Srb10p only becomes evident when Fcp1p activity is reduced. Since inactivation of Srb10p only partially suppresses the phenotypes and the accumulation of hyperphosphorylated CTD caused by mutating fcpl, another kinase must also phosphorylate the CTD in a manner similar to Srb10p. Consistent with this, we find that the fcpl-1 srb10D290A strain contains mostly hyperphosphorylated Rpb1p when grown at the non-permissive temperature of 37°C (M.S.K. and J. Greenblatt, unpublished data). Fig. 5A summarizes the opposing roles of Fcp1p and Srb10p, proposes the existence of another unidentified CTD kinase (x), and indicates the likely effects of these kinases on transcription initiation. Future genome-wide expression studies in fcpl SRB10 and fcpl srb10D290A strains will be necessary to validate this model.
Induction of the RNR3 gene in response to HU might be particularly sensitive to disturbances of the normal CTD phosphorylation pattern caused by mutating fcp1. RNR3 induction is strongly reduced in the fcp1-l strain but is restored to wild-type levels in the fcp1-l srb10D290A strain and therefore closely resembles the growth phenotypes of these strains in response to HU. Interestingly, both uninduced and HU-induced levels of RNR3-lacZ in the srb10D290A strain are significantly higher than in a wild-type strain. It is possible that RNR3 is a direct target for Srb10p mediated transcriptional repression.

Although our studies are consistent with the RNAPII CTD phosphorylation pattern being the major reason for the sensitivity of fcp1 strains to HU and MMS, it is also possible that Fcp1p plays a more direct role in the expression of genes induced by these drugs. It is notable that Fcp1p contains a BRCT domain that is essential for Fcp1p to function in vivo (Kobor et al. 2000) and is usually found in proteins that are involved in the cell cycle checkpoint control in response to DNA damage (Bork et al. 1997). A strain with a mutation in tfgl, encoding a version of RAP74 with strongly reduced binding to Fcp1p, is somewhat sensitive to MMS but not to HU and has a normal CTD phosphorylation pattern (Kobor et al. 2000). Therefore, it is possible that the detailed mechanisms underlying sensitivity to HU and MMS in fcp1 strains are only partly overlapping.

We also propose that Fcp1p and Srb10p interact functionally in a gene-specific manner at some promoters that are regulated by Srb10p. Invasive growth caused by the srb10D290A allele is greatly reversed in fcp1 srb10D290A double mutants. Preliminary genome-wide expression studies indicate that some of the genes, like RNR3, that are up-
regulated in the \textit{srb10D290A} strain are less up-regulated in the \textit{fcp1-I srb10D290A} strain, but the mRNA levels of other Srb10p-regulated genes do not change (N. J. Krogan, Paul Jorgensen, M.S.K., and J. Greenblatt, unpublished data). The genes oppositely affected by Srb10p and Fcp1p include the \textit{GLK1, HXX1, HXT6} and \textit{GAL5} genes that are involved in sugar metabolism and transport (N. J. Krogan, M.S.K., and J. Greenblatt, unpublished data). Interestingly, carbon metabolism may be connected to invasive growth in haploid cells (Palecek et al. 2000). The expression of most genes does not change in \textit{srb10D290A} strains when wild-type \textit{FCP1} is present (Holstege et al. 1998) and we did not detect a global change in the CTD phosphorylation pattern. Srb10p might normally act only on certain promoters at which Srb10p-mediated CTD phosphorylation leads to repression of transcription (Hengartner et al. 1998). Inactivation of Srb10p then results in reduced CTD phosphorylation and up-regulation of these Srb10p-controlled genes. The partial reversal of this up-regulation for certain genes in the \textit{fcp1-I srb10D290A} strain is presumably due to reduced CTD phosphatase activity. This could only be explained if there is an unidentified kinase (Y) that acts similarly to Srb10p, as shown in Fig. 5B. CTD phosphorylation and repression would be partly restored by kinase Y in \textit{fcp1 srb10D290A} strains.

The synthetic phenotypes of \textit{kin28-16 fcp1} strains are in striking contrast to the mutual suppression of phenotypes in the \textit{fcp1 srb10D290A} strains. Concomitant reduction of CTD kinase and CTD phosphatase activity would result in balanced CTD phosphorylation and mutual suppression. Therefore, our data suggests that Kin28p and Fcp1p act at different stages of the transcription cycle or have different substrate
specificities. For example, Fcp1p may act primarily prior to the formation of an initiation complex, whereas TFIIH acts only after transcriptional initiation. Alternatively, whereas TFIIH may preferentially phosphorylate Ser5 of the YPSTSPS heptapeptide repeats within the 5' regions of genes (Komarnitsky et al. 2000), it is not yet clear whether Fcp1p is active in vivo in dephosphorylating Ser5, as opposed to Ser2. To resolve this issue, more information is required about the role of Fcp1p in dephosphorylating Ser2 and Ser5 during chain elongation by RNAPII.
Figure 5. Models for the roles of Fcp1p, Srb10p and another CTD kinase in regulating CTD phosphorylation

A) Model for the involvement of Fcp1p, Srb10p and an unidentified CTD kinase in recycling RNAPII. CTD phosphorylation levels and effects on transcription initiation in three different situations are diagrammed. The sizes of the arrows reflect the level of activity of either Fcp1p (CTD dephosphorylation) or Srb10p and Kinase X (CTD phosphorylation). The predicted relative amounts of RNAPIIA and RNAPIIO in the yeast cells of the different genotypes are indicated by the number of “+” symbols.

B) Model for the functional interaction of Fcp1p, Srb10p and an unidentified CTD kinase on promoters of certain genes regulated by Srb10p. This model is based on the previously suggested model for Srb10p-mediated repression of transcription (Hengartner et al. 1998). The size of the arrows corresponds to the level of activity of the phosphatase Fcp1p or the kinase Srb10p and Kinase Y. It is possible that the Kinase X and Kinase Y are the same protein.
Materials and Methods

Plasmids

To study in vivo phenotypes of kin28-16 fcp1 double mutants, plasmids were constructed by excising KpnI/SacI fragments from plasmids pFK1 [pRS314-FCPI], pFK4 [pRS314-fcp1-1], pFK7 [pRS314-fcp1-2] and pLS2 [pRS314-fcp1-4] (Kobor et al. 1999; Kobor et al. 2000) and inserting them into pASZ11 [ADE2 CEN/ARS] (Stotz and Linder 1990) cut with KpnI and SacI to give pMK104 [pASZ11-FCPI], pMK105 [pASZ11-fcp1-1], pMK106 [pASZ11-fcp1-2], and pMK107 [pASZ11-fcp1-4], respectively.

Yeast strain construction

All strains are derivatives of W303-1A(ade2-1 can1-100 trpl-1 leu2-3.112 his3-11.15 ura3-1 ssd1-d2). Strain YMK197 (kin28::HIS3, p[kin28-16 TRP CEN/ARS], fcp1::LEU2, p[FCPI URA3 CEN/ARS] was obtained by mating yDO4 (kin28::HIS3, p[kin28-16 TRP CEN/ARS], generously provided by Opher Gileadi) with YMK16 (Kobor et al. 1999). Diploids were selected on synthetic complete (SC) plates lacking Trp, His, Leu, and Ura. The haploid strain YMK197 was obtained after sporulation and hydrophobic spore enrichment (Ausubel 1997) and is a kin28-16 strain that contains a chromosomal FCP1 deletion and is kept alive by pMK86 (pRS316-FCPI). YMK16, which was made TRP+ by introducing pRS314, and YMK197 were used for plasmid shuffling experiments with plasmids pMK104, pMK105, pMK106 and pMK107. The transformed cells were grown at 22°C for 4 days on SC plates lacking Ade, Leu, Trp and
His and then streaked on SC plates containing 5-fluoroorotic acid at 22°C for 4 days to counterselect the URA3 marker.

Strains carrying the srb10D290A allele at the chromosomal locus were obtained by cutting plasmid pSK90, an integrating plasmid carrying the URA3 marker (Kuchin and Carlson 1998), with HpaI and introducing this linearized plasmid into YMK18 (Kobor et al. 1999) using the lithium acetate procedure. Transformants were selected on SC plates lacking Ura, Trp, and Leu. Integration was confirmed by testing these strains for invasive growth and also by PCR. The srb10D290A cells were then grown on SC plates lacking Leu and subsequently plated on 5-fluoroorotic acid to select for cells that had lost the URA3 marker. A second test for invasiveness was performed to confirm that the srb10D290A allele was still present in the cells that had lost the URA3 marker. A positive strain carrying the srb10D290A allele was then used to eliminate the plasmid pFK1 (pRS314-FCPI) and substitute pMK86 (pRS316-FCPI) (Kobor et al. 1999) by repeated streaking on SC plates lacking Trp. This gave YMK161, which is a srb10D290A strain with a chromosomal FCPI deletion that is kept alive by pMK86. Plasmids pFK1, pFK4, pFK7 and pLS2 (Kobor et al. 1999; Kobor et al. 2000) were then introduced into the YMK161 strain and transformants were selected on SC plates lacking Trp and Leu.

Plasmid shuffling experiments to counterselect the URA3 marker were performed as above to give strains YMK162 (FCPI, srb10D290A), YMK164 (fcp1-l, srb10D290A), YMK166 (fcp1-2, srb10D290A) and YMK168 (fcp1-4, srb10D290A).

Yeast cell growth assays
The *fcp1 srb10D290A* and *fcp1 kin28-16* double mutant strains were spotted on synthetic complete plates lacking the appropriate amino acids at various temperatures and compared to *fcp1* mutant strains carrying wild type alleles of *SRB10* and *KIN28*. Sensitivity to hydroxyurea (HU) and methanosulfonic methyl ester (MMS) was assessed at 30°C on SC plates lacking the appropriate amino acids and supplemented with either 100 mM HU or 0.0075 % MMS. Invasive growth was assayed by growing yeast strains for 5 days on YPD plates and then washing the plates under running tap water.

**Western blot analysis**

Extracts from yeast strains grown at 30°C in synthetic complete medium lacking the appropriate amino acids were prepared by glass bead lysis in the presence of trichloroacetic acid essentially as described previously (Kobor et al. 1999). After SDS-PAGE and transfer to a nitrocellulose membrane, Western blot analysis using standard procedures was performed using the monoclonal antibody G2 (generously provided by V. Svetlov and R. Burgess) directed against a conserved domain within the largest subunit of yeast RNAPII, a monoclonal antibody against the Rpb3p subunit of yeast RNAPII (Neoclone), an affinity-purified polyclonal antibody against yeast Fcp1p (Kobor et al. 1999) or a monoclonal antibody directed against the vacuolar H⁺ ATPase Vma1p (Molecular Probes).

**β-galactosidase assays**
Liquid culture assays to measure β-galactosidase activity produced by yeast strains containing the RNR3-lacZ reporter plasmid pZZ13 (URA3 CEN/ARS; generously provided by S. Elledge) (Zhou and Elledge 1992) were performed by using cells permeabilized with Sarkosyl as described previously (Kobor et al. 2000). Cells were grown in SC medium lacking Trp, Leu, and Ura in the presence or absence of 100 mM HU for 4 hours at 30°C. For each measurement, β-galactosidase activity was determined from at least three independent cultures and average values are given.
Chapter 5

Reduced Activity of the RNA polymerase II CTD Phosphatase Fcp1p Leads to Increased Synthesis of the Largest Subunit of RNA polymerase II in *Saccharomyces cerevisiae*

I performed the experiments in Fig. 1A, 1B, 2, 3A and 4 and made all the yeast strains used in this Chapter. David Jansma performed the experiments in Fig. 1C and 3B.
Abstract

In this study we show that reduced activity of the yeast RNA polymerase II carboxy-terminal domain phosphatase Fcp1p leads to increased synthesis of the largest subunit of the polymerase. This is mediated by nucleic acid elements upstream of the translation start codon of the gene encoding the largest subunit of RNA polymerase II. The additional largest subunit polypeptides are not extensively phosphorylated on their carboxy-terminal domains (CTDs) and not assembled into RNAPII enzyme complexes. Instead, they accumulate as free subunits at 30°C or are mostly degraded at 37°C. Depletion of Fcp1p therefore results in the accumulation of assembled RNA polymerase II molecules with mostly hyperphosphorylated CTDs. Yeast cells apparently respond in a futile way to reduced CTD phosphatase activity by attempting to overproduce unphosphorylated RNAPII molecules that are competent for transcriptional initiation.
Introduction

Regulated transcriptional initiation by RNA polymerase II (RNAP II) requires the assembly of a multiprotein complex containing RNAP II, the general transcription factors and the SRB/mediator complex involved in both positive and negative control of mRNA synthesis (Lee and Young 2000). The largest subunit of RNAP II contains a unique carboxy-terminal domain (CTD) that consists of tandem heptapeptide repeats with the consensus sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser, the number of which varies among species from 52 repeats in man to 26 or 27 repeats in S. cerevisiae (Allison et al. 1985; Corden et al. 1985). The CTD apparently has important roles in transcription initiation, promoter clearance, chain elongation and transcript processing (Hirose and Manley 2000; Proudfoot 2000).

Whereas the CTD in the transcription initiation complex is hypophosphorylated, it becomes heavily phosphorylated after transcriptional initiation and appears to remain in the hyperphosphorylated form during chain elongation (Cadena and Dahmus 1987). Consistent with this, the CTDs of yeast RNAP II holoenzymes which contain the SRB/mediator complex and are competent for transcription initiation are hypophosphorylated (Kim et al. 1994; Koleske and Young 1994). In contrast, the yeast elongator complex, believed to have a role in regulating transcription elongation, is associated with a form of RNAP II that contains a hyperphosphorylated CTD (Otero et al. 1999).

The known protein kinases that can phosphorylate the CTD at various stages of the transcription cycle include the general transcription factor TFIH, yeast Srb10p and
its human counterpart CDK8, human P-TEFb and yeast CTDK-1 (Lee and Young 2000).

A CTD phosphatase, FCP1, that dephosphorylates the CTD \textit{in vitro} was purified from HeLa cells and \textit{S. cerevisiae} (Chambers and Dahmus 1994; Chambers and Kane 1996). Yeast \textit{FCP1} is essential for viability and encodes a protein with the phosphatase catalytic motif \(\Psi\Psi\Psi\Psi DXDX(T/V)\Psi\Psi\) (\(\Psi\)=hydrophobic residue) that is important for Fcp1p function \textit{in vivo} and \textit{in vitro} (Archambault et al. 1997; Kobor et al. 1999). This motif is contained in the conserved amino-terminal FCP homology domain (FCPH), which resembles similar domains in a number of other proteins (Archambault et al. 1997). These include Psr1p, which is also a protein phosphatase (Siniossoglou et al. 2000). Thus, Fcp1p is the founding member of a new family of protein phosphatases with the characteristic motif \(\Psi\Psi\Psi\Psi DXDX(T/V)\Psi\Psi\) previously identified in a large family of small molecule phosphotransferases and phosphohydrolases (Collet et al. 1998). Recombinant human and yeast Fcp1p are able to dephosphorylate the CTD and the artificial substrate p-nitrophenylphosphate \textit{in vitro} (Cho et al. 1999; Kobor et al. 1999). Fcp1p is also required to dephosphorylate the CTD and for mRNA synthesis in yeast (Kobor et al. 1999; Kobor et al. 2000). Human CTD phosphatase has also been shown to have a role in transcription elongation and in recycling of RNAPII \textit{in vitro} (Cho et al. 1999; Lehman and Dahmus 2000).

In this study, we present evidence for a regulatory mechanism that leads to increased synthesis of the largest subunit of RNAPII when CTD phosphatase activity is reduced. This is mediated mostly by sequences upstream of the translation start codon of
the *RPR1* gene, which encodes the largest subunit of RNAPII. We show that depletion of Fcp1p *in vivo* leads to increases in both the hypophosphorylated and hyperphosphorylated forms of Rpb1p. In these circumstances, however, only the hyperphosphorylated form of Rpb1p is assembled into RNAPII enzymes, whereas the additional, newly synthesized, hypophosphorylated Rpb1p molecules accumulate as unassembled moieties in cell extracts. These orphan subunits are relatively stable at 30°C but are degraded at 37°C.
Results

Destruction of Fcp1p leads to increased levels of Rpb1p

We used a yeast strain in which the pre-existing Fcp1p could be destroyed and new synthesis of Fcp1p could be prevented in order to investigate the effect in vivo of the CTD phosphatase Fcp1p on Rpb1p, the largest subunit of RNAPII. In this strain, the repressible ANBI promoter regulates the expression of a URLF-FCPI fusion gene, which is integrated into the chromosome and encodes ubiquitin-tagged Fcp1p (hereafter referred to as the ANB-URLF-FCPI strain). In this system (Moqtaderi et al. 1996), CuSO4 is used to induce expression of ROXI, whose product represses the ANBI promoter, and UBRI, whose product aids in the destruction of the existing ubiquitin-Fcp1p fusion protein. The ANB-URLF-FCPI strain did not grow on plates containing 500 μM CuSO4, whereas it did grow in the absence of CuSO4 at a rate comparable to that of a control strain carrying wild-type FCPI (Fig. 1A). This is consistent with the essential function of FCPI in vivo (Archambault et al. 1997).

We next examined the phosphorylation state of the Rpb1p in the ANB-URLF-FCPI strain. Whole-cell extracts were prepared at various times after the addition of 500 μM CuSO4 to the growth medium. As expected, the levels of Fcp1p decreased rapidly and significantly in these extracts, although a small amount of partly proteolyzed Fcp1p remained (Fig. 1B). Since the hypophosphorylated IIa form of Rpb1p migrates faster in an SDS polyacrylamide gel than the hyperphosphorylated IIo form, we could compare the amounts of RNAPIIa and RNAPIIo by western blotting these extracts with a
Figure 1. Depletion of Fcp1p leads to an increased level of Rpb1p

A) Depletion of Fcp1p arrests yeast growth. A yeast strain carrying the \textit{ANB-URLF-FCPl} fusion gene and an isogenic control strain carrying a \textit{TRPl CEN/ARS} plasmid were grown at 30°C on synthetic complete solid medium lacking Trp and containing 500 μM CuSO₄. Control plates did not contain any CuSO₄.

B) Depletion of Fcp1p causes a specific increase in the amount of Rpb1p. Whole-cell extracts were prepared from the \textit{FCPl} and \textit{ANB-URLF-FCPl} strains shown in Fig. 1A at various times after the addition of 500 μM CuSO₄ to the growth medium. These extracts were western blotted with the monoclonal antibodies G2, which is directed against a conserved region of Rpb1p outside of the CTD, and 8WG16, which, under these conditions, predominantly recognizes hypophosphorylated CTD. The same blots were also probed with antibodies against Rpb3p, Fcp1p, and the vacuolar ATPase Vma1p.

C) Depletion of Fcp1p leads to increased expression of an \textit{RPBI-lacZ} reporter gene. β-galactosidase activity was measured in the \textit{FCPl} and \textit{ANB-URLF-FCPl} strains carrying a \textit{RPBI-lacZ} reporter gene before and 5 hours after the addition of 500 μM CuSO₄ to the growth medium.
A

B

C

**A**

FCP1 URLF-FCP1

- CuSO₄  + CuSO₄

**B**

Time after Cu²⁺ addition (h)

FCP1 URLF-FCP1

<table>
<thead>
<tr>
<th>0</th>
<th>2</th>
<th>4</th>
<th>6</th>
<th>0</th>
<th>2</th>
<th>4</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- RNAPIIlo
- RNAPIla
- RNAPIIa
- Rpb3p
- Fcp1p
- Vma1p

**C**

β-Galactosidase activity (Miller units)

<table>
<thead>
<tr>
<th>0</th>
<th>200</th>
<th>400</th>
<th>600</th>
<th>800</th>
<th>1000</th>
<th>1200</th>
<th>1400</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- minus Cu
- plus Cu

FCP1 URLF-FCP1
monoclonal antibody G2 which recognizes an epitope on Rpb1p outside the CTD. We reasoned that destroying CTD phosphatase in cells should lead to a shift in the ratio between RNAPIIa and RNAPIIo. In control cells, we observed no significant change in the ratio of hypophosphorylated to hyperphosphorylated Rpb1p (Fig. 1B). In extracts prepared from the \textit{ANB-URLF-FCP1} strain a dramatic increase in both forms of the largest subunit of RNAPII occurred as judged by western blotting with the G2 antibody. However, visual inspection of the western blot did not indicate a major change in the ratio between the two forms of the largest subunit. These results were confirmed by western blots with the 8WG16 antibody, which under the conditions used, recognizes mostly the hypophosphorylated form of the CTD (Fig. 1B). In contrast, the amount of the Rpb3p subunit of RNAPII did not change significantly (Fig. 1B). Also, western blotting revealed no significant change in the vacuolar H$^+$-ATPase Vma1p, which was used as a gel-loading control. These results suggested that yeast cells react to depletion of the CTD phosphatase through a mechanism that increased the amount of Rpb1p.

The nature of this effect was investigated by using a reporter construct that allowed us to measure \textit{RPBI} expression. The reporter construct, a fusion of \textit{RPBI} in frame with the bacterial \textit{lacZ} gene on a high-copy 2-$\mu$ plasmid included the upstream \textit{SCM3} ORF, the \textit{RPBI} upstream and core promoter elements, the 5'-untranslated and the first 171 nucleotides of the \textit{RPBI} ORF. We found that depletion of Fcp1p caused an approximately fivefold increase in $\beta$-galactosidase activity (Fig. 1C). In contrast, addition of CuSO$_4$ did not lead to an increase of $\beta$-galactosidase activity in the \textit{FCPI} control
238

strain. Therefore, the initial portion of the *RPBI* ORF and the region upstream of the *RPBI* ATG codon are sufficient to mediate the increased synthesis of Rpb1p in response to a depletion of Fcp1p. These results were consistent with our findings from the western blot experiments (Fig. 1B).

**Depletion of Fcp1p leads to an increased proportion of assembled RNAPII molecules containing a hyperphosphorylated CTD**

Our results indicated that Rpb1p accumulates in yeast cells when Fcp1p activity is reduced, whereas the amount of the Rpb3p subunit of RNAPII does not change. The extracts that we used in these experiments were prepared by glass-bead lysis in the presence of trichloroacetic acid, and therefore represent denatured and precipitated total cellular protein. To determine how this increased amount of Rpb1p was reflected in assembled RNAPII enzyme molecules, we constructed an *ANB-URLF-FCP1* strain in which 13 Myc epitopes were fused to the carboxy-terminus of Rpb3p. We then used a monoclonal antibody that recognizes the Myc epitope to immunoprecipitate RNAPII from the soluble fraction of yeast whole-cell extracts. Prior to the addition of CuSO₄, most of the immunoprecipitated RNAPII contained hypophosphorylated Rpb1p, whereas the RNAPII present in the soluble fraction after the depletion of Fcp1p was almost exclusively in the hyperphosphorylated RNAPII form (Fig. 2, lanes 3 and 4). In order to see Rpb1p and Rpb3p on the same gel, conditions were used in this experiment that do not allow for clear separation of the two forms of Rpb1p. The additional hypophosphorylated Rpb1p polypeptides that remained after the depletion of Fcp1p
Figure 2. Excess Rpb1p subunit is not assembled into RNAPII

RNAPII was immunoprecipitated by using anti-Myc antibody and Rpb3p tagged at its carboxy-terminus with 13 Myc epitopes. Pellet and supernatant fractions from the immunoprecipitations from the soluble fraction of yeast extracts prepared either before or 6 hours after the addition of 500 μM CuSO₄ to the growth media were western blotted with the monoclonal antibody G2 and the Rpb3p antibody.
500μM CuSO₄

<table>
<thead>
<tr>
<th></th>
<th>No Tag</th>
<th>Rpb3p-Myc Tag</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pellet</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rpb1p</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rpb3p-Myc</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Supernatant</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Rpb1p</td>
<td>Rpb3p</td>
</tr>
</tbody>
</table>
were found in the supernatant after the immunoprecipitation (Fig. 2, lane 8). No detectable hypophosphorylated Rpb1p was present in the supernatant after immunoprecipitation of extracts prepared before the addition of CuSO₄ to the culture, indicating that the immunoprecipitation with anti-Myc antibody was efficient (Fig. 2, lane 7). The exclusive presence of Myc-tagged Rpb3p subunit in the pellet fraction also demonstrated the efficiency of the immunoprecipitation. We concluded that the accumulated hypophosphorylated Rpb1p produced in response to Fcp1p depletion is not incorporated into functional enzyme but exists as free polypeptides in yeast cells. These data also showed for the first time that depletion of Fcp1p results in the exclusive presence of assembled hyperphosphorylated RNAPII in the soluble fraction of yeast cell extracts.

Degradation of free Rpb1p at an elevated temperature

In the course of our investigations with strains carrying various fcp1 mutant alleles we observed increased amounts of Rpb1p in cells grown at 30°C (see Fig. 4B), whereas a shift of cells to 37°C led to the complete disappearance of hypophosphorylated Rpb1p (Kobor et al. 1999; Kobor et al. 2000). These observations prompted us to investigate the fate of the free hypophosphorylated Rpb1p polypeptides in the ANB-URLF-FCPI strain at an elevated temperature. We used a protocol in which either we shifted cells directly to 37°C after the addition of 500 μM CuSO₄ or first depleted Fcp1p for 3 hours at 30°C and then shifted the cells to 37°C for an additional 3 hours. Extracts prepared from these cells
Figure 3. Free Rpb1p is degraded at 37°C

A) Western blot analysis of whole-cell extracts prepared from FCP1 and ANB-URLF-FCPI yeast strains grown at either 30°C or 37°C, as indicated, at various times after the addition of CuSO4. An arrow indicates that this sample was shifted to 37°C for an additional 3 hours after 3 hours of growth at 30°C. The same panel of antibodies was used as in the experiment described in Fig. 1B. Short and long exposures are presented for the blot with G2 monoclonal antibody.

B) The compensatory response on RPB1 expression is intact at 37°C. β-galactosidase activity produced from the RPB1-lacZ reporter plasmid was measured after growth for the indicated times at the indicated temperatures.
**A**

<table>
<thead>
<tr>
<th>Temp. (°C)</th>
<th>Time (h)</th>
<th>8WG16</th>
<th>FCP1</th>
<th>URLF-FCP1</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- FCP1
- URLF-FCP1

**B**

<table>
<thead>
<tr>
<th>Time</th>
<th>Temp</th>
<th>CuSO₄</th>
<th>FCP1</th>
<th>URLF-FCP1</th>
<th>Fold Increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 h</td>
<td>30°C</td>
<td>-</td>
<td>162+/−22</td>
<td>195+/−67</td>
<td>1.2</td>
</tr>
<tr>
<td>3 h</td>
<td>30°C</td>
<td>+</td>
<td>165+/−64</td>
<td>450+/−110</td>
<td>2.7</td>
</tr>
<tr>
<td>6 h</td>
<td>30°C</td>
<td>+</td>
<td>207+/−9</td>
<td>686+/−92</td>
<td>3.3</td>
</tr>
<tr>
<td>6 h</td>
<td>30°C</td>
<td>+</td>
<td>186+/−5</td>
<td>814+/−84</td>
<td>4.4</td>
</tr>
<tr>
<td>6 h</td>
<td>37°C</td>
<td>+</td>
<td>169+/−31</td>
<td>523+/−9</td>
<td>3.1</td>
</tr>
</tbody>
</table>
0, 3 and 6 hours after the addition of CuSO₄ were then examined by western blotting for the relative amounts of Rpb1p in the IIa and IIo forms. As expected, cells from the ANB-URLF-FCP1 strain grown for 3 hours at 30°C in the presence of 500 µM CuSO₄ contained a much higher amount of Rpb1p than cells from the congenic control strain (Fig. 3A). However, when the former cells were shifted to 37°C, the amount of hypophosphorylated Rpb1p decreased, while the amount of Rpb1b in the hyperphosphorylated form was not affected (Fig. 3A). Similarly, the ANB-URLF-FCP1 strain shifted to 37°C immediately following the addition of 500 µM CuSO₄ contained mostly hyperphosphorylated Rpb1p. The amount of hyperphosphorylated Rpb1p in the ANB-URLF-FCP after shift of the cells 37°C appeared higher than the total amount of Rpb1p in the control cells, indicating that some hyperphosphorylated Rpb1p might not be incorporated into RNAPII enzyme. The control strain that lacked the ANB-URLF-FCP1 fusion gene contained an unchanged level of Rpb1p. The amounts of Rpb3p and Vma1p remained relatively constant, whereas Fcp1p was destroyed only in the ANB-URLF-FCP1 strain (Fig. 3A).

We used the RBP1-lacZ reporter construct to test whether the disappearance of the hypophosphorylated Rpb1p at 37°C correlated with a decrease in the expression of Rpb1p. β-galactosidase activity was measured in the ANB-URLF-FCP1 strain and the control strain at the same times and temperatures as in the western blot experiment (Fig. 3B). As before, we observed an increase of β-galactosidase activity in the ANB-URLF-FCP1 strain in the presence of 500 µM CuSO₄. The β-galactosidase activity increased by
2.7-fold 3 hours after the addition of CuSO$_4$ and by 3.3-fold after 6 hours (Fig. 3B). When the cells were shifted to 37°C for 3 hours, after having been grown for 3 hours at 30°C, we detected a 4.4-fold increase in $\beta$-galactosidase activity. Also, the ANB-URLF-FCPI strain shifted to 37°C for 6 hours immediately following the addition of CuSO$_4$ showed a 3.1-fold increase in $\beta$-galactosidase activity. These data suggested that the response of $RPBl$ expression to a depletion of Fcp1p is still active at 37°C and that the disappearance at 37°C of the free Rpb1p is likely caused by protein degradation.

Reduced compensatory response in cells with a replacement of the $RPBl$ regulatory regions

We showed previously that the depletion of Fcp1p in a strain in which the upstream region of $RPBl$ is replaced by that of LEU2 leads to an increased amount of Rpb1p with a hyperphosphorylated CTD and a decreased amount of Rpb1p with a hypophosphorylated CTD (Kobor et al. 1999). This indicated that Fcp1p is required for dephosphorylation of the CTD in vivo and also implicated sequences upstream of the $RPBl$ ATG codon as being important for the observed increase in the amount of this protein (Kobor et al. 1999). We used either the natural $RPBl$ promoter or the LEU2 promoter to regulate the expression of the $RPBl$ ORF (Fig. 4A) and examined Rpb1p levels and phosphorylation patterns in fcp1 strains with mutations in the FCPH domain (Kobor et al. 1999) in order to further test the involvement of $RPBl$ upstream regions in this response. The LEU2 promoter is well suited as a replacement of the $RPBl$ upstream
Figure 4. Regions upstream of the translation start codon of *RPB1* are involved in the increased synthesis of Rpb1p in response to mutations in *FCP1*

A) Diagram outlining the regulatory region of *RPB1* and the *LEU2* replacement used in Fig. 4B. The UAS and T-rich regions are involved in the regulation of *RPB1* expression (Jansma et al. 1996).

B) Whole cell extracts prepared from *FCP1, fcp1-1* and *fcp1-2* strains with either the natural *RPB1* regulatory region or the *LEU2* regulatory region and grown at 30°C were western blotted with the antibodies described in the legend to Fig. 1B.
region for these experiments since the expression of the *pLEU2-RPBI* fusion gene is
similar to the expression of *RPBI* (Fig. 4B) (Archambault et al. 1996).

The *fcp1* mutant strains with the natural *RPBI* regulatory region contained
elevated levels of both the hypophosphorylated and the hyperphosphorylated forms of
Rpb1p at the permissive temperature, consistent with our previous observations (Kobor
et al. 1999; Kobor et al. 2000). Replacement of the *RPBI* regulatory regions with the
*LEU2* promoter resulted in a significant decrease in the amount of the
hypophosphorylated Rpb1p as judged by western blotting of extracts from these strains
with the 8WG16 antibody, which detects primarily hypophosphorylated Rpb1p (Fig.
4B). Western blotting with the G2 antibody demonstrated a similar lack of increase for
the hypophosphorylated form of Rpb1p, although slight variations in the levels of the
different forms of Rpb1p occurred in this experiment. The amounts of Rpb3p, Fcp1p,
and Vma1p did not change significantly in extracts prepared from these strains (Fig. 4B).
These results indicate that decreased CTD phosphatase activity caused by mutations in
the *FCP1* gene also leads to a higher amount of hypophosphorylated Rpb1p via a
mechanism that is at least partly mediated by regions upstream of the *RPBI* ATG
translational start codon.
Discussion

We have shown that reduced amounts of the RNAPII CTD phosphatase Fcp1p lead to increased levels of both the hypophosphorylated and hyperphosphorylated forms of the largest subunit of RNAPII. Our results indicate that this elevated amount of Rpb1p is produced by yeast cells as a result of a mechanism that is at least partly mediated by the promoter and/or 5'-untranslated region of RPBI. The additional Rpb1p is not assembled into functional RNAPII, but rather exists as free polypeptides that are largely degraded at an elevated temperature. These studies provide clear evidence that depletion of Fcp1p leads to a change in the phosphorylation pattern of RNAPII molecules in vivo. Whereas our previous studies demonstrated a disappearance of hypophosphorylated Rpb1p and accumulation of hyperphosphorylated Rpb1p in yeast extracts (Kobor et al. 1999; Kobor et al. 2000), our current results show that lack of CTD phosphatase always leads to loss of the hypophosphorylated IIa subunit and accumulation of the hyperphosphorylated IIo subunit in assembled RNAPII enzyme molecules. Because the loss of Fcp1p leads to a shutdown in transcription by RNAPII (Kobor et al. 1999; Kobor et al. 2000), these results suggest that a major essential function of Fcp1p is the creation of initiation-competent RNAPIIA during the recycling of the transcription machinery. This is consistent with recent in vitro studies (Cho et al. 1999).

Regulatory mechanisms that maintain an optimal amount of functional RNAPII have been identified in Escherichia coli, Caenorhabditis elegans and mammalian cells (Somers et al. 1975; Crerar and Pearson 1977; Guialis et al. 1977; Dennis and Fill 1979; Dennis et al. 1985; Downing and Dennis 1991; Steward and Linn 1992; Dalley et al. 1993;
Dykxhoorn et al. 1996). Hamster cells that contain both wild-type and α-amanitin-resistant forms of RPB1 show an increased rate of synthesis of the largest subunit by up to 2.3-fold when grown in the presence of α-amanitin (Guialis et al. 1979), a toxin that inhibits RNAPII and causes selective degradation of the α-amanitin-sensitive (i.e. wild-type) RPB1 subunit (Nguyen et al. 1996). Experimental conditions that promote cardiac hypertrophy, such as α1-adrenergic stimulation and aortic banding, also lead to an increase in both forms of the largest subunit of RNAPII, although it is not known whether the excess RBP1 protein is actually assembled into RNAPII in these cases (Abdellatif et al. 1998). Our studies suggest that in yeast cells a reduction in CTD phosphatase activity leads to elevated synthesis of Rpb1p. This mechanism may be an emergency measure that results in the preservation of the maximum possible amount of initiation-competent RNAPIIA.

The increased synthesis of Rpb1p in response to reduced Fcp1p activity is likely mediated by a post-transcriptional rather than a transcriptional mechanism. Using DNA microarrays we did not observe any increase in RPBI mRNA levels in fcp1-1 mutant cells grown at the permissive temperature (N.J. Krogan and J. Greenblatt, unpublished data), even though the cells show a significant increase in Rpb1p (Fig. 4B, lanes 1-3). One possibility is that yeast cells contain genes that are particularly sensitive to disturbances in the CTD phosphorylation pattern or, more generally, in the general RNAPII transcription machinery. Some genes, such as INO1 (Archambault et al. 1996), are very sensitive to a reduced activity of the transcription machinery. A similar gene may exist
that encodes an inhibitor of the translation of \textit{RPB1} mRNA (Fig. 5). Compromising the expression of such a gene by depleting Fcp1p could lead to reduced levels of the inhibitor, followed by increased translation of \textit{RPB1} mRNA and consequently higher levels of Rpb1p (Fig. 5). Our studies provide evidence for involvement of the 5'-untranslated region of \textit{RPB1} in this regulation. Dissecting this region will give further insight into the nucleic acid elements that are involved.
Figure 5. Model for the regulation of the translation of the *RPBI* mRNA in response to reduced Fcp1p activity
We also found that the excess unassembled Rpb1p that is produced in response to reduced Fcp1p activity is stable at 30°C but disappears almost completely at 37°C. Since the increased synthesis of Rpb1p remains when the cells are grown at 37°C, the disappearance of Rpb1p at 37°C likely is due to protein degradation. We propose the existence of a cellular mechanism for degrading unassembled, hypophosphorylated Rpb1p at this temperature. Several components of the cellular protein-degradation machinery have been linked to RNAPII. The ubiquitin ligase Rsp5p can bind to the CTD of Rpb1p and plays a role in Rpb1p degradation (Huibregtse et al. 1997; Chang et al. 2000). Mutations in rpb1 can be suppressed by mutations in kex2, which encodes an endoprotease located in the Golgi apparatus (Martin and Young 1989). The Pub3p subunit of the 26S proteasome plays a role in the degradation of Rpb6p, a subunit common to all three nuclear RNA polymerases (Nouraini et al. 1997). Whether the degradation of free Rpb1p reported here is mediated by any of these proteins remains subject to further investigation.
Materials and Methods

Plasmids

The reporter plasmid that uses β-galactosidase activity to assess the functioning of the RPBI regulatory region (pDJ80) was constructed by inserting a HindIII-BamHI fragment containing RPBI sequences from −1583 to +171 (the A of the initiation codon is +1) into the HindIII and BamHI sites of YEp356R (Myers et al. 1986).

Plasmid pMK56 was obtained in a two-step procedure. A 1000 bp fragment of FCP1 starting at the ATG was PCR amplified and inserted into the EcoRI and XbaI sites of ZMY168 (Moqtaderi et al. 1996), thereby creating a plasmid containing the ANB-URLF-FCPI cassette. The cassette was then excised by cutting this plasmid with KpnI and SacI and subsequently inserted into the KpnI and SacI sites of the integrating vector pRS304 (Sikorski and Hieter 1989), which carries the TRPI marker for selection in yeast.

Yeast strains

Yeast strains YMK18 (FCPI), YMK20 (fcpl-1) and YMK26 (fcpl-2) are derivatives of W303-1A containing the natural RPBI regulatory elements. YMK36 (FCPI), YMK38 (fcpl-1) and YMK40 (fcpl-2) carry the same FCP1 alleles but the regions upstream of the ATG of RPBI were replaced by the LEU2 regulatory regions. Details of the strain construction have been described previously (Kobor et al. 1999).

YMK2 is a derivative of ZMY60 (Moqtaderi et al. 1996) in which a fusion gene of the ANB-URLF cassette with FCP1 was generated by introducing the integrating
plasmid pMK56 (previously linearized with BstBI, which cuts at codon 136 of FCPI) and selecting for Trp\(^+\) colonies. A short, nonfunctional 5'-fragment of the original FCPI is still present in this strain. YMK1 is an isogenic control strain carrying a TRP1 CEN/ARS plasmid instead of the ANB-URLF-FCPI fusion gene.

A linear DNA fragment encoding 13 Myc epitopes and the kanamycin-resistance gene was obtained by PCR using primer MK174 RPB3 Tag F2 [5'-GCA TCT CAA ATG GGT AAT ACT GGA TCA GGA GGG TAT GAT AAT GCT TGG CGG ATC CCC GGG TTA ATT AA-3'] and primer MK175 RPB3 Tag R1 [5'-TTA TTT TCG GTT CGT TCA CTT GTT TTT TTT CCT CTA TTA CGC CCA CTT GAG AAT TCG AGC TCG TTT AAA C-3'] with pFA6a-13Myc-kanMX6 (Longtine et al. 1998) as a template. This fragment contains 50 nucleotides of homology to the 3'-end of the RPB3 gene and 50 nucleotides of homology to regions downstream of the stop codon of RPB3. It was introduced into YMK2 using the Li-Acetate procedure and transformants were selected on YPD agar medium containing 200 \(\mu\)g/ml G418. Integration of this construct at the natural RPB3 locus was confirmed by western blotting using a monoclonal anti-Myc antibody (a gift from J. Moffat and B. Andrews) and by PCR analysis. The strain was designated YMK251.

\[ \text{\textbeta-galactosidase assays} \]

\[ \text{\textbeta-galactosidase activity was measured as described previously (Jansma et al. 1996).} \]

Cells were grown in synthetic complete medium lacking the appropriate amino acids at
30°C or 37°C. For each measurement, β-galactosidase activity was determined in at least three independent cultures.

Western Blot Analysis

Protein extracts were prepared as described previously from yeast cells growing logarithmically in synthetic complete medium lacking the appropriate amino acids (Kobor et al. 1999). Western analysis was performed by standard procedures using monoclonal antibody G2 directed against a conserved epitope within Rpb1p (generously provided by V. Svetlov and R. Burgess), the 8WG16 monoclonal antibody directed primarily against hypophosphorylated CTD, a monoclonal antibody against Rpb3p (Neoclone), an-affinity purified polyclonal antibody against yeast Fcp1p (Kobor et al. 1999), or a monoclonal antibody against the vacuolar H⁺-ATPase Vma1p (Molecular Probes).

Immunoprecipitation of RNAPII

Yeast strains YMK2 (ANB-URLF-FCPI) and YMK251 (ANB-URLF-FCPI; RPB3-13Myc), 50 A₆₀₀ units of each, were collected by centrifugation, washed once with ice-cold water and frozen at −70°C. The cells were resuspended in 500 μl of IP-lysis buffer (50 mM Tris pH7.5, 250 mM NaCl, 0.1% NP40, 10 mM sodium pyrophosphate, 5 mM EGTA, 5 mM EDTA, 0.1 mM orthovanadate, 5 mM NaF, 1 mM dithiothreitol and protease inhibitors (Boehringer Mannheim, Complete™) and broken by glass-bead lysis by vortexing 8×30 sec at 4°C. Insoluble material was removed by centrifugation. 50
μl of the soluble yeast extract was diluted tenfold in IP-lysis buffer and incubated with 5 μg of anti-Myc monoclonal antibody and 20 μl of protein A sepharose (Pharmacia) for 2 hours at 4°C. The beads were pelleted by centrifugation at low speed and washed three times with 1 ml of IP-lysis buffer. Finally, beads were boiled in SDS-PAGE sample buffer and analyzed by western blotting.
Chapter 6

Summary and Perspectives
The work that I present in this thesis establishes Fcp1p as a RNAPII CTD phosphatase, examines effects of Fcp1p on gene expression and provides insights into various aspects of its regulation. My research on Fcp1p's enzymatic activity, its protein-protein interactions, and the genetic interactions involving the *FCP1* gene is summarized in Figure 1.
Figure 1. Summary of genetic interactions, physical interactions and catalytic activity involving *FCPI* described in this thesis.

Genetic interactions are diagrammed by dashed arrows and protein-protein interactions by continuous arrows. See text for further details.
I showed in Chapter 2 that Fcp1p is an unusual eukaryotic protein phosphatase that contains the YΨΨΨDXDX(T/V)ΨΨ motif previously identified in the catalytic centers of small molecule phosphatases and phosphotransferases. This motif is essential for Fcp1p to function in vivo and essential for Fcp1p to dephosphorylate both the CTD and the chromogenic artificial substrate p-nitrophenylphosphate in vitro. The ΨΨΨDXDX(T/V)ΨΨ motif is contained within an larger FCPH domain, which is homologous to similar domains found in a number of proteins of unknown function, including the putative tumor suppressor protein HYA22 (Ishikawa et al. 1997) and the putative oncoprotein OS4 protein (Su et al. 1997). Since the work in Chapter 2 was published, Psr1p, which also contains an FCPH domain, has been shown to be a protein phosphatase and to have a ΨΨΨDXDX(T/V)ΨΨ motif that is important for catalysis (Siniossoglou et al. 2000). Therefore, I postulate that Fcp1p is the founding member of a family of eukaryotic protein phosphatases which is distinct from the families described so far. To confirm this, other members of the FCPH family should be tested for protein phosphatase activity. Chapter 2 also shows that Fcp1p is also required for dephosphorylation of the CTD in vivo, and is essential for the expression of most genes in yeast.

The work presented in Chapter 3 shows that the BRCT domain of Fcp1p is essential for Fcp1p to function in vivo. The interaction of Fcp1p with the general transcription factors TFIIF and TFIIB is also described in Chapter 3. A cDNA encoding
human FCP1a was originally recovered in a two-hybrid screen using human RAP74 as a bait (Archambault et al. 1998b). Moreover, the interaction of yeast Fcp1p with yeast RAP74 was also described previously (Archambault et al. 1997). The work in Chapter 3 provides a more detailed mapping of the regions of Fcp1p that mediate its binding to RAP74. Two independent portions in the carboxy-terminal half of Fcp1p are shown to bind RAP74.

Conversely, I showed in Chapter 3 that a KEFGK motif in RAP74 is involved in its binding to Fcp1p. I also discovered a similar motif in TFIIB and showed that TFIIB and RAP74 interact in similar ways with Fcp1p \textit{in vitro}. The KEFGK motif in TFIIB is contained within the first cyclin-related repeat of TFIIB, and an amino acid substitution in this motif reduces its affinity for Fcp1p. Similarly, amino acid changes in the KEFGK motif of RAP74 reduce binding to Fcp1p. I also showed in Chapter 3 that yeast strains with the \textit{tfgl-1} mutation, encoding RAP74 with amino acid changes in its KEFGK motif, have synthetic phenotypes when \textit{fcp1} is also mutated, providing evidence for the physiological importance of the RAP74-Fcp1p interaction. Interestingly, yeast strains that have Fcp1p with a carboxy-terminal truncation are viable despite the failure of this version of Fcp1p to bind to both RAP74 and TFIIB \textit{in vitro}. It is curious that a similar \textit{fcp1} allele cannot support yeast viability when the gene encoding the putative transcription elongation factor Rtf1p is mutated (Costa and Arndt 2000). Although this carboxy-terminal region of Fcp1p and its amino-terminal region are not essential for viability, a form of the gene expressing Fcp1p that lacks both termini cannot support viability.
I also showed in Chapter 3 that Fcp1p can activate transcription when artificially tethered to a promoter by fusing it to the DNA-binding domain of the bacterial LexA protein. Fcp1p-mediated activation depends on its carboxy-terminal domain that interacts with TFIIF. Consistent with this, activation by LexA-Fcp1p is strongly reduced in a tfgl-2 mutant strain, suggesting that this activation involves binding of Fcp1p to RAP74, which can, in turn, interact with RNAPII and presumably recruit it to a promoter.

In Chapter 4, I described work aimed at understanding the biological interactions between Fcp1p and the two known CTD kinases found in RNAPII holoenzymes, Srb10p and Kin28p. Yeast strains with mutations in fcp1 have much more severe growth defects when kin28, encoding the kinase subunit of TFIIH, is also mutated. This genetic interaction is consistent with a role for Fcp1p in the recycling of RNAPII after transcription terminated and suggests that Fcp1p and TFIIH are active at different stages of the transcription cycle.

In striking contrast, the temperature-sensitivity of fcp1 strains and the sensitivity of fcp1 strains to certain cell cycle checkpoint inducing drugs are partly suppressed by inactivation of Srb10p. Inactivation of Srb10p also partly reverses the abnormal CTD phosphorylation pattern and higher amount of Rpb1p that I detected in fcp1 strains. Therefore, I propose that Fcp1p and Srb10p have opposing roles in CTD phosphorylation and, likely, in gene expression.

A role for Fcp1p in creating initiation competent RNAPII is further supported by the work that is described in Chapter 5. There, I showed for the first time that depletion of Fcp1p results in the exclusive presence of RNAPIIO and no assembled RNAPIIA in
the soluble fraction of yeast extracts. Reduced Fcp1p activity also leads to increased synthesis of Rpb1p as a consequence of a compensatory mechanism that is partly mediated by nucleic acid sequences upstream of the translational initiation codon of the \textit{RPBI} mRNA. However, the increased amount of Rpb1p is not assembled into RNAPII complexes and is largely degraded at 37°C. The work in Chapter 5 therefore rationalizes my observation that depletion or inactivation of Fcp1p results in an increase in the amount of hyperphosphorylated RNAPII that is partly masked in whole-cell extracts by the excess hypophosphorylated Rpb1p.

Although these studies show that Fcp1p is a CTD phosphatase and greatly enhances our understanding of Fcp1p function in yeast, much work remains to be done, and I will outline some of it below. It is intriguing that Fcp1p is a processive phosphatase (Chambers and Dahmus 1994) but the detailed catalytic mechanism of the Fcp1p phosphatase is not yet clear. Although the data presented in this thesis suggest that the $\Psi^3\Psi^2\Psi^1DXDX(T/V)\Psi^4$ motif is involved in the phosphatase reaction, other residues might also contribute. A more saturating mutagenesis of conserved residues within the FCPH domain, combined with assessing the altered proteins for function \textit{in vivo} and activity \textit{in vitro}, could provide further insight. Also, substrate trapping experiments similar to those that have been described for the small molecule phosphatases and phosphotransferases might identify the exact roles of particular amino acid residues (Collet et al. 1998). In this respect, it has not yet been determined whether the altered versions of Fcp1p encoded by the \textit{fcp1-1}, \textit{fcp1-2} and \textit{fcp1-4} alleles described in this thesis
have reduced phosphatase activity as the in vivo studies may suggest. A high resolution structure of the FCPH domain together with a phosphorylated CTD substrate would probably provide the best possible insight into the catalytic mechanism, and efforts to determine this structure are currently underway in collaboration with the laboratory of Dr. Griesinger (Goettingen, Germany).

It is also unclear whether Fcp1p is equally active on RNAPII when its CTD is phosphorylated with different kinases, an issue that is related to the in vivo specificity of Fcp1 (see below). The experiments described in Chapter 2 used TFIIH to phosphophorylate the CTD. Recombinant or highly purified forms of Srb10p and CTDK-I that could be used to phosphorylate the CTD are also now available. A detailed quantitative comparison should reveal any substrate preferences that Fcp1p might display. The activity of Fcp1p on RNAPII phosphorylated with TFIIH is relatively low, and it is also possible that a co-factor is needed to enhance catalysis. It has been suggested that yeast Fcp1p needs another fraction in order to function in vitro (Chambers and Kane 1996) and, although our studies presented in Chapter 2 disprove this, it is still possible that this fraction is needed for optimal activity. Alternatively, it has been suggested that isomerization of the phosphorylated CTD by the peptidylprolyl-cis-trans-isomerase Ess1p might provide a better substrate for Fcp1p (Wu et al. 2000). This could be tested by first incubating RNAPII phosphorylated in vitro by various CTD kinases with Ess1p followed by a Fcp1p-dependent dephosphorylation reaction.

It is also possible to use an in vivo approach to examine which residues in the CTD are preferentially dephosphorylated by Fcp1p. Extracts from the various fcp
strains described in this thesis can be tested with antibodies specific for either phospho-Ser2 or phospho-Ser5. Because of the difficulties in quantitation due to the increased levels of RNAPIIa described in Chapter 5, it also might be a good idea to examine the assembled RNAPIIO present in the soluble fraction of yeast cell extracts with these antibodies. In addition, RNAPII molecules bound to chromatin in wild-type and fcp1 strains can be crosslinked to DNA and immunoprecipitated with these antibodies. The subsequent quantification of the crosslinked DNA by PCR will provide a measure of the amounts of the various phospho-isoforms of RNAPIIO associated with different regions of a gene. Some of these experiments are currently being performed in collaboration with the laboratory of Dr. S. Buratowski (Harvard Medical School). It is possible that Fcp1p in vivo preferentially dephosphorylates either Ser2 or Ser5. This issue could be genetically dissected by examining yeast strains in which mutations in fcp1 are combined with forms of rpb1 that encode Rpb1p with mutations in either Ser2 or Ser5 of the CTD (West and Corden 1995). It should be cautioned, however, that all these in vivo studies are indirect assays of substrate specificity, in contrast to more direct in vitro studies.

Another major avenue of future investigation is the regulation of Fcp1p activity through its interactions with various proteins. I already showed that both RAP74 and TFIIB can bind Fcp1p and that this binding involves a KEGFK motif common to these factors. The binding of RAP74 and TFIIB is consistent with previous findings that RAP74 can stimulate human CTD phosphatase and this stimulation can be inhibited by TFIIB (Chambers et al. 1995). The same assays can be performed using recombinant or highly purified yeast factors, and then the activities in vitro of either wild-type or
mutated versions of RAP74 and TFIIIB that are described in this thesis can also be tested. I have not yet been able to determine whether RAP74, TFIIIB and Fcp1p form a trimeric complex or, alternatively, whether RAP74 and TFIIIB compete for binding to Fcp1p. The major binding sites for RAP74 and TFIIIB in the carboxy-terminal region of Fcp1p differ slightly, but further investigation will be necessary. Identification of the exact molecular interfaces among these factors is currently being pursued through NMR structure studies using recombinant portions of RAP74, Fcp1p, and TFIIIB in collaboration with Dr. J. Omichinski (University of Georgia). It is possible that Fcp1p interacts with proteins other than RAP74 and TFIIIB. Further interactions might be identified by affinity chromatography using recombinant Fcp1, by two-hybrid screen using a form of the LexA-Fcp1 fusion protein that does not activate on its own (as described in Chapter 3), or by the purification from yeast extract of epitope tagged-versions of Fcp1p.

It is also possible that Fcp1p might be a homodimer, as has been suggested for human FCP (Cho et al. 1999). In this respect, it is interesting that I showed that the BRCT domain of Fcp1p is essential in vivo. BRCT domains frequently are involved in homodimer and heterodimer formation, and it may be that mutations in the BRCT domain impair function by disrupting homodimer formation. This could be tested by comparing recombinant wild-type and mutated forms of Fcp1p in sedimentation experiments. Alternatively, it is possible that the Fcp1p BRCT domain is involved in another essential interaction, e.g. with RNAPII, or has an as yet undiscovered function.

A related issue is the question of what protein complexes contain Fcp1p. LexA-Fcp1p can activate transcription when tethered to a promoter, an activity that has been
proposed for components of the RNAPII holoenzyme. However, I was not able to detect Fcp1p in an RNAPII holoenzyme preparation containing general transcription factors and some Srb proteins (Hirst et al. 1999) when an extract from wild-type cells was used. However, it is possible that Fcp1p only weakly or transiently associates with the RNAPII holoenzyme, which would be consistent with the data that I present in Chapter 4. Human FCP1a seems to be a component of a similar form of the human RNAPII holoenzyme (Archambault et al. 1998b).

In contrast, preliminary data suggest that some fraction of the cellular Fcp1p associates with elongating RNAPIIO (J. Svjestrup, Imperial Cancer Research Fund, unpublished data). Consistent with this, I have found that Fcp1p can bind the elongator subunit Elp1p in vitro and that FCP1 and the ELP genes have a strong genetic interaction (M. S. K. and J. Greenblatt, unpublished data). The presence of Fcp1p in elongation complexes would be consistent with a role for Fcp1p in elongation, as has been proposed as a result of various studies involving Fcp1p (Cho et al. 1999; Costa and Arndt 2000; Lehman and Dahmus 2000). It remains to be seen, however, whether or not the elongation activity of Fcp1p is independent of its CTD phosphatase. It is also entirely possible that Fcp1p has functions and substrates outside of the RNAPII transcription complex that remain to be discovered.

In my opinion, a major issue when exactly in the transcription cycle dephosphorylation of the CTD by Fcp1p takes place and this is related to the signals that regulate Fcp1p activity. Although data presented in this thesis demonstrate that lack of Fcp1p activity leads to the elimination of assembled RNAPIIA and the exclusive
presence of RNAPIIO in the soluble fraction of yeast extracts, it is not clear whether
dephosphorylation occurs before or after transcription termination and how or whether
dephosphorylation relates to release of the RNAPII from the template. This could be
addressed by chromatin immunoprecipitation experiments with wild type and \textit{fcp1} strains
in which phospho-CTD specific antibodies are used to determine whether the CTD is
still phosphorylated after the RNAPII has passed a termination site.

Furthermore, it has also been suggested that the CTD is differentially
phosphorylated by TFIIH and other CTD kinases during the transcription cycle. TFIIH-
dependent phosphorylation seem to be confined to promoter regions and disappears
during early transcription elongation (Komarnitsky et al. 2000; Schroeder et al. 2000). It
will be crucial to determine whether this CTD dephosphorylation is performed by Fcp1p
although the negative genetic interaction between \textit{FCP1} and \textit{KIN28}, presented in Chapter
4 of this thesis, might suggest otherwise. Chromatin immunoprecipitation experiments
using strains described in this thesis are currently being done in collaboration with
Dr.S.Buratowski to investigate these issues.

A lot of the questions regarding Fcp1 function also await the development of an \textit{in}
vitro transcription system that reliably reconstitutes initiation, elongation and
termination, by RNAPII on chromatin templates. Of course, further insight into Fcp1p
function can also be obtained using genetic approaches. In this regard, both suppressor
screens and synthetic lethal screens could be performed using as a starting point the \textit{fcp1}
mutant yeast strains that I described in this thesis.
References


Archambault, J., R.S. Chambers, M.S. Kobor, Y. Ho, M. Cartier, D. Bolotin, B.


Blau, J., H. Xiao, S. McCracken, P. O'Hare, J. Greenblatt and D. Bentley. 1996. Three


phosphatase functions to recycle RNA polymerase II. *Genes Dev* 13: 1540-52.


Dubois, M.F., M. Vincent, M. Vigneron, J. Adamczewski, J.M. Egly and O. Bensaude.


Fondell, J.D., M. Guermah, S. Malik and R.G. Roeder. 1999. Thyroid hormone receptor-associated proteins and general positive cofactors mediate thyroid hormone receptor function in the absence of the TATA box-binding protein-associated factors of TFIID.
Proc Natl Acad Sci U S A 96: 1959-64.


RAP74 induces promoter contacts by RNA polymerase II upstream and downstream of a DNA bend centered on the TATA box. Proc Natl Acad Sci U S A 94: 7150-5.


Kanazawa, S., T. Okamoto and B.M. Peterlin. 2000. Tat competes with CIITA for the binding to P-TEFb and blocks the expression of MHC class II genes in HIV infection. *Immunity* **12**: 61-70.


Keaveney, M. and K. Struhl. 1998. Activator-mediated recruitment of the RNA polymerase II machinery is the predominant mechanism for transcriptional activation in


Kuchin, S. and M. Carlson. 1998. Functional relationships of Srb10-Srb11 kinase,
carboxy-terminal domain kinase CTDK-I, and transcriptional corepressor Ssn6-Tup1.


polymerase II. *Mol Cell Biol* **18**: 2130-42.


carboxy-terminal domain contributes to the response to multiple acidic activators in vitro.


Moyle, M., J.S. Lee, W.F. Anderson and C.J. Ingles. 1989. The C-terminal domain of the


Transcriptional activation by Gcn4p involves independent interactions with the SWI/SNF complex and the SRB/mediator. *Mol Cell* 4: 657-64.


Otero, G., J. Fellows, Y. Li, T. de Bizemont, A.M. Dirac, C.M. Gustafsson, H.


Prelich, G. 1997. *Saccharomyces cerevisiae* BUR6 encodes a DRAP1/NC2alpha homolog that has both positive and negative roles in transcription *in vivo*. *Mol Cell Biol* 17: 2057-


Soulier, J. and N.F. Lowndes. 1999. The BRCT domain of the S. cerevisiae checkpoint


II CTD deletion and substitution mutations. *Genetics* **140**: 1223-33.


specific and phosphorylation-dependent proline isomerization: a potential mitotic

Yamaguchi, Y., T. Takagi, T. Wada, K. Yano, A. Furuya, S. Sugimoto, J. Hasegawa and
H. Handa. 1999a. NELF, a multisubunit complex containing RD, cooperates with DSIF to

Structure and function of the human transcription elongation factor DSIF. J Biol Chem
274: 8085-92.

transcription factor IIIF in promoter escape by RNA polymerase II. J Biol Chem 274:
35668-75.

Yang, L., H.A. Chansky and D.D. Hickstein. 2000. EWS-Fli-1 Fusion Protein Interacts
with Hyperphosphorylated RNA Polymerase II and Interferes with Serine-Arginine

Yang, X., M.O. Gold, D.N. Tang, D.E. Lewis, E. Aguilar-Cordova, A.P. Rice and C.H.
Herrmann. 1997. TAK, an HIV Tat-associated kinase, is a member of the cyclin-dependent family of protein kinases and is induced by activation of peripheral blood lymphocytes and differentiation of promonocytic cell lines. *Proc Natl Acad Sci U S A* 94: 12331-6.


Yuan, C.X., M. Ito, J.D. Fondell, Z.Y. Fu and R.G. Roeder. 1998. The TRAP220 component of a thyroid hormone receptor-associated protein (TRAP) coactivator


Zandomeni, R.O. 1989. Kinetics of inhibition by 5,6-dichloro-1-beta-D-
ribofuranosylbenzimidazole on calf thymus casein kinase II. *Biochem J* **262**: 469-73.


### Plasmids used in this study

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Plasmid Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pMK56</td>
<td>pRS304-ANBI-URLF-FCPI</td>
<td>Chapter 5</td>
</tr>
<tr>
<td>pMK58</td>
<td>pRS314-fcp1 (D180E)</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>pMK59</td>
<td>PRS314-fcp1 (D180N)</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>pMK61</td>
<td>PRS314-fcp1 (D182E)</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>pMK63</td>
<td>PRS314-fcp1 (D182N)</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>pMK65</td>
<td>pGEX-3X-yRAP74 (649-735) K695E</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>pMK66</td>
<td>pGEX-3X-yRAP74 (649-735) K695E/K699E</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>pMK67</td>
<td>pGEX-3X-yRAP74 (649-735) K695A/K699A</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>pMK68</td>
<td>pGEX-3X-yRAP74 (649-735) E696K</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>JA815</td>
<td>pEG202-Fcp1(1-732)</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>JA811</td>
<td>pEG202-Gal4 (pSH17-4)</td>
<td>Roger Brent</td>
</tr>
<tr>
<td>JA821</td>
<td>pEG202-Stop</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>pMK71</td>
<td>pRS306-LEU2 promoter-RPBI ORF (to SpeI site)</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>pMK72</td>
<td>PEG202-Fcp1 (1-666)</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>pMK73</td>
<td>PEG202-Fcp1 (1-644)</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>pMK74</td>
<td>PEG202-Fcp1 (1-626)</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>pMK75</td>
<td>PEG202-Fcp1 (457-732)</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>pMK76</td>
<td>PEG202-Fcp1 (457-666)</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>pMK77</td>
<td>PEG202-Fcp1 (626-732)</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>pMK78</td>
<td>PEG202-Fcp1 (645-732)</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>pMK79</td>
<td>PEG202-Fcp1 (667-732)</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>pMK86</td>
<td>PRS316-FCPI</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>pMK87</td>
<td>pBac-Fcp1</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>pMK88</td>
<td>pBac-Fcp1 (D180E)</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>pMK89</td>
<td>pBac-Fcp1 (D182E)</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>pMK90</td>
<td>pRS314-fcp1-3 (Δ2-134)</td>
<td>Chapter 2</td>
</tr>
</tbody>
</table>
pMK91  pGEX4T-1-TFIIB (1-120)  Chapter 3
pMK92  pGEX4T-1-TFIIB (100-345)  Chapter 3
pMK93  pGEX4T-1-TFIIB (100-240)  Chapter 3
pMK94  pGEX4T-1-TFIIB (200-345)  Chapter 3
pMK95  pGEX4T-1-TFIIB (100-345) K201E  Chapter 3
pMK96  pRS314-fcp l-5 (1-594)  Chapter 3
pMK97  pRS314-fcp l-6 (1-557)  Chapter 3
pMK98  pRS314-fcp l-7 (135-594)  Chapter 3
pMK99  pRS303-TRG1 (636-735) HA  Chapter 3
pMK100  pRS303-fgg l-2 (636-735) HA K695E/K699E  Chapter 3
pMK101  pRS314-fcp l-8 (1-626)  Chapter 3
pJA785  pET23d-HA-Fcp l (1-666)  Archambault et al 1997
pMK102  pET23d-HA-Fcp l (1-594)  Chapter 3
pMK103  pET23d-HA-Fcp l (1-626)  Chapter 3
pMK104  pASZ11-FCP l  Chapter 4
pMK105  pASZ11-fcp l-1 (R250A/P251A)  Chapter 4
pMK106  pASZ11-fcp l-2 (L177A/L181A/H187A)  Chapter 4
pMK107  pASZ11-fcp l-4 (W575A)  Chapter 4
PET32a  Vector for Thioredoxin (TRX) fusion proteins  Novagen
pFK1  pRS314-FCPl  Chapter 2
pFK4  pRS314-fcp l-1 (R250A/P251A)  Chapter 2
pFK7  pRS314-fcp l-2 (L177A/L181A/H187A)  Chapter 2
pLS2  pRS314-fcp l-4 (W575A)  Chapter 3
pRB1840  LacZ reporter with 1 LexA binding site  Origene
pJK103  LacZ reporter with 2 LexA binding sites  Origene
pSH18-34 \textit{LacZ} reporter with 8 LexA binding sites

pDJ80 \textit{pRBPI-LacZ}

Origene

Chapter 5
Yeast strains used in this study

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant genotype</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>YMK1</td>
<td>FCP1, p[TRP1 CEN/ARS]</td>
<td>Chapter 5</td>
</tr>
<tr>
<td>YMK2</td>
<td>fcp1::TRP1, ANB1-URLF-FCPI</td>
<td>Chapter 5</td>
</tr>
<tr>
<td>YMK16</td>
<td>MATα.fcp1Δ::LEU2, p[FCPI URA3 CEN/ARS]</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>YMK18</td>
<td>MATα.fcp1Δ::LEU2, p[FCPI TRP1 CEN/ARS]</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>YMK20</td>
<td>MATα.fcp1Δ::LEU2, p[fcp1-1 TRP1 CEN/ARS]</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>YMK26</td>
<td>MATα.fcp1Δ::LEU2, p[fcp1-2 TRP1 CEN/ARS]</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>YMK28</td>
<td>MATα.fcp1Δ::LEU2, p[fcp1-3 TRP1 CEN/ARS]</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>YMK32</td>
<td>MATα.fcp1Δ::LEU2 pLEU2-RPB1 leu2-3, 112 or LEU2 p[FCPI URA3 CEN/ARS]</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>YMK34</td>
<td>MATα.fcp1Δ::LEU2 pLEU2-RPB1 leu2-3, 112 or LEU2 p[FCPI TRP1 CEN/ARS]</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>YMK36</td>
<td>MATα.fcp1Δ::LEU2 pLEU2-RPB1 leu2-3, 112 or LEU2 p[fcp1-1 TRP1 CEN/ARS]</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>YMK38</td>
<td>MATα.fcp1Δ::LEU2 pLEU2-RPB1 leu2-3, 112 or LEU2 p[fcp1-2 TRP1 CEN/ARS]</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>YMK40</td>
<td>MATα.fcp1Δ::LEU2 pLEU2-RPB1 leu2-3, 112 or LEU2 p[fcp1-3 TRP1 CEN/ARS]</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>YMK88</td>
<td>FCP1, p[TRP1 CEN/ARS], pLEU2-RPB1</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>YMK89</td>
<td>fcp1::TRP1, ANB1-URLF-FCPI, pLEU2-RPB1</td>
<td>Chapter 2</td>
</tr>
<tr>
<td>YMK110</td>
<td>MATα.fcp1Δ::LEU2, p[fcp1-4 TRP1 CEN/ARS]</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>YMK161</td>
<td>MATα.fcp1Δ::LEU2, p[FCPI URA3 CEN/ARS], srb10D290A</td>
<td>Chapter 4</td>
</tr>
<tr>
<td>YMK162</td>
<td>MATα.fcp1Δ::LEU2, p[FCPI TRP1 CEN/ARS], srb10D290A</td>
<td>Chapter 4</td>
</tr>
<tr>
<td>YMK164</td>
<td>MATα.fcp1Δ::LEU2, p[fcp1-1 TRP1 CEN/ARS], srb10D290A</td>
<td>Chapter 4</td>
</tr>
<tr>
<td>YMK166</td>
<td>MATα.fcp1Δ::LEU2, p[fcp1-2 TRP1 CEN/ARS], srb10D290A</td>
<td>Chapter 4</td>
</tr>
<tr>
<td>YMK168</td>
<td>MATα.fcp1Δ::LEU2, p[fcp1-4 TRP1 CEN/ARS], srb10D290A</td>
<td>Chapter 4</td>
</tr>
<tr>
<td>YMK197</td>
<td>MATα.kin28::HIS3, p[kin28-16 TRP1 CEN/ARS], fcp1Δ::LEU2, p[FCPI URA3 CEN/ARS]</td>
<td>Chapter 4</td>
</tr>
<tr>
<td>YMK202</td>
<td>MATα.tfg1::TFG1::HIS3 FCP1</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>YMK204</td>
<td>MATα.tfg1::tfg1-2::HIS3 FCP1</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>YMK211</td>
<td>MATα.tfg1::TFG1::his3::TRP1 FCP1</td>
<td>Chapter 3</td>
</tr>
</tbody>
</table>
YMK212  MATα tfgl::tfgl-2::his3::TRP1 FCP1  Chapter 3
YMK215  MATα tfgl::TFG1::HIS3 fcp1Δ::LEU2, p[FCP1 URA3 CEN/ARS]  Chapter 3
YMK217  MATα tfgl::tfgl-2::HIS3 fcp1Δ::LEU2, p[FCP1 URA3 CEN/ARS]  Chapter 3
YMK219  MATα fcp1Δ::LEU2, p[FCP1 ADE2 CEN/ARS], p[TRPI  Chapter 4
CEN/ARS]
YMK220  MATα fcp1Δ::LEU2, p[fcp1-1 ADE2 CEN/ARS], p[TRPI  Chapter 4
CEN/ARS]
YMK221  MATα fcp1Δ::LEU2, p[fcp1-2 ADE2 CEN/ARS], p[TRPI  Chapter 4
CEN/ARS]
YMK222  MATα fcp1Δ::LEU2, p[fcp1-4 ADE2 CEN/ARS], p[TRPI  Chapter 4
CEN/ARS]
YMK223  MATα kin28::HIS3, p[kin28-16 TRP1 CEN/ARS]. fcp1Δ::LEU2,  Chapter 4
p[FCP1 ADE2 CEN/ARS]
YMK224  MATα kin28::HIS3, p[kin28-16 TRP1 CEN/ARS], fcp1Δ::LEU2,  Chapter 4
p[fcp1-1 ADE2 CEN/ARS]
YMK225  MATα kin28::HIS3, p[kin28-16 TRP1 CEN/ARS], fcp1Δ::LEU2,  Chapter 4
p[fcp1-2 ADE2 CEN/ARS]
YMK226  MATα kin28::HIS3, p[kin28-16 TRP1 CEN/ARS], fcp1Δ::LEU2,  Chapter 4
p[fcp1-4 ADE2 CEN/ARS]
YMK227  MATα tfgl::TFG1::HIS3 fcp1Δ::LEU2, p[FCP1 TRP1 CEN/ARS]  Chapter 3
YMK228  MATα tfgl::TFG1::HIS3 fcp1Δ::LEU2, p[fcp1-1 TRP1 CEN/ARS]  Chapter 3
YMK229  MATα tfgl::TFG1::HIS3 fcp1Δ::LEU2, p[fcp1-1 TRP1 CEN/ARS]  Chapter 3
YMK230  MATα tfgl::TFG1::HIS3 fcp1Δ::LEU2, p[fcp1-2 TRP1 CEN/ARS]  Chapter 3
YMK231  MATα tfgl::TFG1::HIS3 fcp1Δ::LEU2, p[fcp1-3 TRP1 CEN/ARS]  Chapter 3
YMK232  MATα tfgl::TFG1::HIS3 fcp1Δ::LEU2, p[fcp1-4 TRP1 CEN/ARS]  Chapter 3
YMK233  MATα tfgl::tfgl-2::HIS3 fcp1Δ::LEU2, p[FCP1 TRP1 CEN/ARS]  Chapter 3
YMK234  MATα tfgl::tfgl-2::HIS3 fcp1Δ::LEU2, p[fcp1-1 TRP1 CEN/ARS]  Chapter 3
YMK235  MATα tfgl::tfgl-2::HIS3 fcp1Δ::LEU2, p[fcp1-2 TRP1 CEN/ARS]  Chapter 3
YMK236  MATα tfgl::tfgl-2::HIS3 fcp1Δ::LEU2, p[fcp1-3 TRP1 CEN/ARS]  Chapter 3
YMK237  \( \text{MAT}^\alpha \text{tfg}^{1}:\text{tfg}^{1-2}:\text{HIS}3 \text{fcp}^{1}\Delta:\text{LEU}2, p[\text{fcp}^{1-4} \text{TRP}^1 \text{CEN/ARS}] \)  Chapter 3

YMK238  \( \text{MAT}^\alpha \text{tfg}^{1}:\text{tfg}^{1-2}:\text{HIS}3 \text{fcp}^{1}\Delta:\text{LEU}2, p[\text{fcp}^{1-5} \text{TRP}^1 \text{CEN/ARS}] \)  Chapter 3

YMK251  \( \text{MAT}^\alpha \text{fcp}^{1}:\text{TRP}^1, \text{ANB}1-\text{URLF-FCP}^1, \text{rpb}3:\text{rpb}3-\text{MYC}13: \)  Chapter 5

\( \text{Kan}^{**} \)