The function of the protein phosphatase Glc7p in transcription termination, RNA processing and transcriptional regulation of ribosomal protein genes

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Salome Röck
aus Deutschland

Basel, 2007

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von
Prof. Walter Keller, Dr. Helge Grosshans und Prof. Markus Affolter

Basel, den 27.03.07

Prof. Hans-Peter Hauri
Dekan der Philosophisch-
Naturwissenschaftlichen Fakultät

Table of contents

Summary

Gene transcription in general can be subdivided into three main phases: transcription initiation, elongation and termination. The enzyme that accomplishes transcription of protein coding genes, snRNAs and snoRNAs is RNA polymerase II (RNAP II). During transcription, the nascent RNA is processed in several ways in order to generate a mature functional RNA. For this, the transcripts of protein coding genes are capped at the 5' end, introns are spliced out and the 3' ends are processed by endonucleolytic cleavage at the poly(A) site followed by poly(A) tail synthesis (polyadenylation). In yeast, the cleavage and polyadenylation reaction requires a 3' end processing complex consisting of the cleavage and polyadenylation factor (CPF), cleavage factor IA (CF IA), cleavage factor IB (CF IB) and the poly(A) binding protein. In contrast to pre-mRNAs, most pre-snoRNAs are processed only at their 3' end. Furthermore, snoRNAs are not polyadenylated.

CPF is not only involved in 3' end processing, but distinct subunits of CPF have additional functions in transcription elongation and termination of mRNAs and snoRNAs.

In recent years affinity purification of the CPF complex has lead to the identification of several new subunits of CPF (Ohnacker et al., 2000). Among them is the essential protein phosphatase Glc7p, the yeast homologue of mammalian protein phosphatase 1 (PP1). Glc7p has diverse cellular functions (Stark, 1996). The specificity of a reaction that requires Glc7p is accomplished by targeting or regulatory factors that direct Glc7p to the location of the reaction or regulate its activity.

The aim of this thesis was to study the function of Glc7p as part of CPF. In Chapter 2 we show, that Glc7p is required for the polyadenylation but not for the cleavage step of pre-mRNA 3' end processing *in vitro* and *in vivo*. In addition, Glc7p is needed for correct poly(A) site selection. Glc7p physically interacts with several subunits of CPF and CF IA. One of them, the CPF subunit Pta1p, has been reported to be dephosphorylated by Glc7p (He and Moore, 2005). Dephosphorylation of Pta1p stimulates the polyadenylation reaction. Thus, Glc7p regulates polyadenylation via the phosphorylation state of Pta1p (He and Moore, 2005). We also observed that in *glc7* mutant strains, several subunits of CPF are underrepresented. This might indicate that the activity of Glc7p is required for the formation of stable CPF complexes.

Analysis of several *glc7* mutants also revealed that Glc7p is involved in transcription termination of snoRNAs (Chapter 3). Our data suggest that Glc7p functions in the Nrd1 complex-dependent pathway of snoRNA transcription termination. However, none of the Nrd1 complex subunits was found to be a target for dephosphorylation by Glc7p. In contrast, Glc7p is not involved in transcription termination of pre-mRNAs.

A reduction in poly(A)-dependent pausing in *glc7* mutants indicated that Glc7p might also be involved in regulating transcription elongation (Chapter 4). Further investigation showed that Glc7p genetically interacts with the transcription elongation factors Spt4p, Leo1p and Rtf1p. In addition, several *glc7* mutants are sensitive to the drug 6-azauracil (6AU). Sensitivity to 6AU is a phenotypic landmark of transcription elongation mutants. Interestingly, the snoRNA transcription termination defect observed in *glc7* mutants is suppressed in *glc7/spt4*, *glc7/leo1* and *glc7/rtf1* double mutants. This suggests that Glc7p acts as a factor required for snoRNA transcription termination that modifies transcription elongation factors to facilitate transcription termination. Therefore, Glc7p might couple transcription elongation to transcription termination.

Microarray analysis of the temperature sensitive *glc7-12* allele (Chapter 5) indicated that Glc7p is involved in transcription regulation of ribosomal protein (RP) and Ribi genes. Two signaling pathways control the transcription of RP and Ribi genes in response to environmental conditions: the target of rapamycin (TOR) and the Ras/PKA signaling pathway. These pathways regulate the localization of the transcription factors Fhl1p, Ifh1p, Crf1p and Sfp1p to RP or Ribi gene promoters. Epistasis experiments suggest that Glc7p acts downstream of the signaling component PKA to regulate the transcription of RP genes (Chapter 6). In addition, we found that Glc7p controls the nuclear localization of Yak1p and Crf1p. Yak1p is a downstream target of the kinase PKA. Crf1p in turn is phosphorylated by Yak1p, shuttles to the nucleus and represses transcription of RP genes. Regulation of the localization of the co-repressor Crf1p by Glc7 could represent one of several redundant ways to suppress transcription of RP genes.

1 Introduction: processing and transcription termination of mRNAs, snoRNAs and snRNAs

1.1 Eukaryotic gene transcription and processing of messenger RNAs

In metazoans, three RNA polymerases (RNAPs) are responsible for the transcription of genes. RNAP I transcribes genes coding for ribosomal RNAs (rRNAs). RNAP III transcribes genes coding for transfer RNAs (tRNA) and in addition the 5S rRNA. RNAP II transcribes protein-coding genes and also genes for non-coding small nuclear RNAs (snRNAs) and small nucleolar RNAs (snoRNAs). In the following introduction I will concentrate on transcripts generated by RNAP II.

The precursors of mRNAs (pre-mRNAs), snoRNAs and snRNAs have to undergo a number of processing events to become mature functional RNAs. Processing of pre-mRNAs includes capping, splicing and 3' end processing (cleavage and polyadenylation), which all occur during the transcription process (co-transcriptionally) in the nucleus of eukaryotic cells. Transcription of genes by RNAP II can be divided into three main phases: initiation, elongation and termination of transcription (Saunders et al., 2006). During transcription initiation RNAP II is recruited to the promoter with the help of general transcription factors and a transcription initiation complex is formed that is associated with the template stably enough to efficiently transcribe the full length of the gene. Transcription of the full length gene is referred to as transcription elongation and will be discussed in more detail below. Transcription terminates when RNAP II dissociates from the DNA template and from the 3'UTR of the RNA transcript.

Soon after transcription has initiated, the nascent pre-mRNA is capped at the 5' end (Shatkin and Manley, 2000). The 7-methylguanine cap is added to the 5' triphosphate end of the pre-mRNA when the transcript is around 20-25 nucleotides long. In *S. cerevisiae* capping requires the action of three capping enzymes: Cet1p, an RNA triphosphatase, which hydrolyzes the 5' triphosphate to a 5' diphosphate; Ceg1p, an RNA guanylyltransferase, which adds GMP to the 5' end; and finally Abd1p a methyltransferase that methylates the GMP at position N7 (Shatkin and Manley, 2000).

Another cotranscriptional event is splicing of pre-mRNAs, which removes introns from the newly transcribed pre-mRNA and joins exons together. A large

complex, the spliceosome, composed of small nuclear ribonucleoprotein particles (snRNPs) and non-snRNP proteins achieves splicing by two transesterification reactions (Kramer, 1996).

3' end processing of pre-mRNAs is a two step reaction. The pre-mRNA is endonucleolyticaly cleaved and subsequently polyadenylated (Wahle and Rüegsegger, 1999.; Figure 1.1). The cleavage and polyadenylation reaction is performed by the cleavage and polyadenylation factor (CPF) complex together with the cleavage factor IA (CF IA). Poly(A) tails are important for mRNA stabilization, mRNA export and also for translation of mRNAs (Long et al., 1995; Sachs and Wahle, 1993). Finally, transcription termination occurs at random positions within the 3'UTR and is coupled to 3' end processing (Hirose and Manley, 2000). During transcription termination RNAP II dissociates from the template DNA and from the nascent RNA and can be recycled to start a new round of transcription.

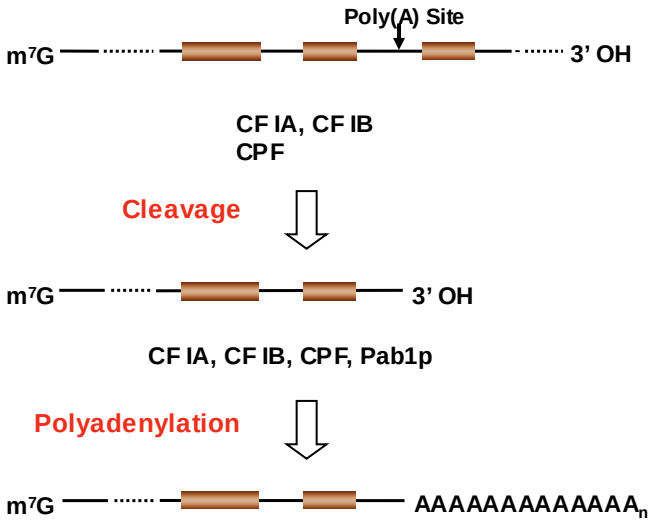


Figure 1.1. The two steps of 3' end processing

The precursor mRNA (top of scheme) is cleaved at the poly(A) site. Factors required for the cleavage step are indicated. The upstream cleavage product is subsequently polyadenylated. Factors required for polyadenylation are indicated.

1.2 3’ end processing of pre-mRNAs

1.2.1 Cis-acting signals are required for 3’ end processing of pre-mRNAs

3’ end processing (endonucleolytic cleavage and subsequent polyadenylation of pre-mRNAs) of pre-mRNA occurs in the 3’ UTRs of pre-mRNAs and is guided by a number of cis-acting signals. In yeast, up to four separate cis-acting elements were found. All cis-acting elements contribute to the correct recognition of the processing site thus ensuring the accuracy and efficiency of cleavage and polyadenylation (Graber et al., 1999a; Graber et al., 1999b). In contrast to higher eukaryotes, cis-acting signals in *S. cerevisiae* are of statistical nature meaning they lack strong consensus sequences.

The four cis-acting signals are: the efficiency element (EE), the positioning element (PE), the poly(A) site and the U-rich element (Graber et al., 1999a; Guo et al., 1995; Wahle and Rüegsegger, 1999; Zhao et al., 1999a; Figure 1.2). The EE is composed of versions of the sequence UAUAUA which is the most efficient version of the EE (Guo et al., 1995; Guo and Sherman, 1996). Moreover, in yeast the EE has been demonstrated to be more important than the PE in selecting the processing site (Guo and Sherman, 1996). EEs appear to be used only rarely in mammals (Graber et al., 1999a).

The PE is an A-rich element, frequently AAUAAA, centered approximately 20 nucleotides upstream of the poly(A) site (Graber et al., 1999b). This element is found in all organisms investigated to date. Whereas the sequence AAUAAA is highly conserved in mammals, the PE in yeast and plants varies significantly (Graber et al., 1999a).

Mapping of the poly(A) site, which is equal to the cleavage site, revealed that polyadenylation occurs most frequently at a $Py(A)_n$ sequence (Py = pyrimidine; Heidmann et al., 1994). In contrast to animal genes, in which a single poly(A) site is found downstream of AAUAAA, in many yeast genes a cluster of poly(A) sites downstream of the EE and PE are found (Zhao et al., 1999a). *In vivo*, mutations in the PE most often do not affect the overall efficiency of processing but instead cause a defocusing of the endonucleolytic activity from the native poly(A) site to cryptic poly(A) sites (Guo and Sherman, 1996).

The poly(A) site is often preceded by U-rich elements in yeast, motifs that do not have counterparts in animals. In addition, the poly(A) site is also commonly followed by a downstream U-rich signal. It was recently reported, that the U-rich signals contribute to the cleavage/poly(A) site selection and enhance 3' end processing efficiency *in vitro* (Dichtl and Keller, 2001).

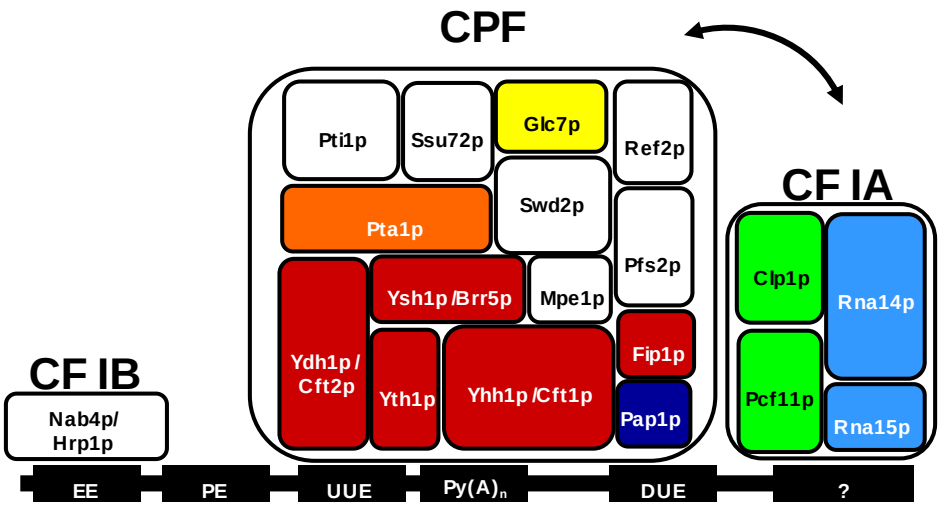


Figure 1.2. Cis- and trans-acting factors required for 3' end processing of pre-mRNAs

Schematic illustration of the polypeptide composition of the trans-acting factors involved in 3' end formation and their relative position to the cis-acting sequenes required for this reaction. EE efficiency element; PE positioning element; UUE upstream U-rich element; $Py(A)_n$ poly(A) site; DUE downstream U-rich element;

1.2.2 Trans-acting factors required for 3' end processing of pre-mRNAs

Cleavage and polyadenylation of pre-mRNAs is a two step reaction. At first view one might consider two enzymes to be sufficient for this reactions: an endonuclease that cleaves the pre-mRNA at the poly(A) site and a poly(A) polymerases (PAP) that synthesizes the poly(A) tail. However, intensive biochemical characterization of yeast and mammalian trans-acting factors uncovered a large number of factors required for 3' end processing of pre-mRNAs (Keller and Minvielle-Sebastia, 1997; Wahle and Rüegsegger, 1999; Zhao et al., 1999a). While the polyadenylation signals used by mammals and yeast are rather different in consensus sequence and organization, the factors which comprise the

cleavage/polyadenylation apparatus in these two organisms are conserved to a high degree (Zhao et al., 1999a).

Fractionation of whole-cell yeast extracts has identified five functionally distinct activities involved in cleavage and polyadenylation (Chen and Moore, 1992; Kessler et al., 1996). Cleavage and polyadenylation factor IA (CF IA), IB (CF IB) and II (CF II) are sufficient for the cleavage reaction, while specific poly(A) addition requires CF IA, CF IB, Pap1p, the poly(A) binding protein (Pab1p) and the polyadenylation factor I (PF I). A multi-protein complex, designated cleavage and polyadenylation factor (CPF), where PF I/Pap1p and CF II activities are combined has recently been isolated from yeast extracts by an one-step affinity purification (Ohnacker et al., 2000). It is now generally accepted that PF I and CF II form a functional unit *in vivo*.

CF IA consists of four polypeptides, Rna14p, Rna15p, Pcf11p and Clp1p (Amrani et al., 1997a; Kessler et al., 1996; Minvielle-Sebastia et al., 1994). Temperature-sensitive *rna14* and *rna15* mutant strains are defective in both cleavage and polyadenylation of a synthetic precursor RNA (Minvielle-Sebastia et al., 1994). In addition, these alleles are synthetic lethal with mutations in the PAP1 gene (Minvielle-Sebastia et al., 1994). Rna15p has a N-terminal RNA-recognition motif (RRM) and has affinity to U-rich sequences (Kessler et al., 1996; Minvielle-Sebastia et al., 1991; Takagaki and Manley, 1997). Recently, it was demonstrated that Rna15p recognizes the PE (Gross and Moore, 2001a). The interaction of Rna15p with PE requires the presence of Rna14p and Nab4p/Hrp1p (Gross and Moore, 2001a). Pcf11p mutant cells display a deficiency in both cleavage and polyadenylation *in vitro* and the poly(A) length is strongly reduced in vivo (Amrani et al., 1997a). Pcf11p has a N-terminal CTD-interaction domain (CID) that is common to CTD-binding proteins like yeast Nrd1p (Barilla et al., 2001; Steinmetz et al., 2001). In addition to its role in 3' end processing, Pcf11p is also involved in transcription termination of pre-mRNAs (Sadowski, 2003). Clp1p contains Walker A and B motifs, which have been implicated in ATP-GTP binding (Preker et al., 1997). Mutant Clp1p extracts are deficient in cleavage and polyadenylation (T. Wiederkehr, unpublished result).

Purified CF IB is a single polypeptide encoded by the NAB4/HRP1 gene (Kessler et al., 1997). This gene was previously identified as a suppressor of a temperature-sensitive *npl3* allele, a gene encoding a protein which is involved in mRNA export (Henry et al., 1996). Nab4p contains two RNA binding domains in its

middle region and RNA-binding and SELEX analyses suggest that it interacts with the EE (Chen and Hyman, 1998; Henry et al., 1996; Valentini et al., 1999). Nab4 is essential for the polyadenylation but is not required for the cleavage step. Furthermore, selection of the correct cleavage site requires Nab4 in a concentration-dependent manner (Minvielle-Sebastia et al., 1998). Synthetic lethality and two-hybrid analyses indicate that Nab4p interacts with Rna14p and Rna15p (Kessler et al., 1997). Interestingly, Nab4p is also required for nonsense-mediated mRNA decay (NMD) in the cytoplasm (Gonzalez et al., 2000). NMD is a surveillance mechanism that monitors premature translation termination and degrades aberrant mRNAs (Hentze and Kulozik, 1999).

Pab1p has four N-terminal RNA-binding domains and associates with the poly(A) tails of mRNAs in both the nucleus and the cytoplasm (Adam et al., 1986; Burd and Dreyfuss, 1994; Swanson and Dreyfuss, 1988). The main functions of Pab1p include the poly(A)-dependent translation initiation (Tarun and Sachs, 1996), the deadenylation-dependent mRNA turnover (Caponigro and Parker, 1996) and the poly(A) tail length control (Minvielle-Sebastia et al., 1997).

CPF is comprised of the subunits Yhh1p, Ydh1p, Ysh1p, Pta1p, Pap1p, Ref2p, Mpe1p, Fip1p, Pfs2p, Pti1p, Glc7p, Swd2p, Yth1p and Ssu72p (Dichtl et al., 2002b). Yhh1p is an RNA-binding protein which is involved in poly(A) site recognition (Dichtl and Keller, 2001). In addition to its requirement in the cleavage and polyadenylation reaction, Yhh1p is involved in transcription termination of pre-mRNAs (Dichtl et al., 2002b). Interestingly, Yhh1p specifically interacts with the phosphorylated CTD. The direct interactions of Yhh1p with nascent RNA and the CTD suggests that it communicates poly(A) site recognition to the elongating RNAP II to initiate transcription termination (Dichtl et al., 2002b).

Mutations in Ydh1p inhibit both the cleavage and the polyadenylation step of pre-mRNA 3' end processing (Kyburz et al., 2003). In addition, RNA-binding experiments showed that Ydh1p binds with high affinity to sequences around the poly(A) site and more weakly to the EE (Dichtl and Keller, 2001). Subsequently, it was shown that *ydh* mutant strains are deficient in poly(A) site recognition (Kyburz et al., 2003) suggesting that binding of Ydh1 to sequences around the poly(A) site is important for poly(A) site selection.

Ysh1p has a β-lactamase motif, which is common to metal-dependent hydrolases (Callebaut et al., 2002). Because of its similarity to other nucleases in the

metallo-beta-lactamase family, the Brr5/Ysh1 subunit has been proposed to be the endonuclease.

Pta1p is exclusively required for the polyadenylation step in pre-mRNA 3' end processing (Preker et al., 1997). The phosphorylation state of Pta1p is regulated by the phosphatase Glc7p and an unknown kinase (He and Moore, 2005). He and co-workers demonstrated that a phosphorylated form of Pta1p inhibits polyadenylation whereas unphosphorylated Pta1p does not (He and Moore, 2005). This suggests that yeast mRNA polyadenylation is regulated by phosphorylation.

Pap1p is the enzyme that synthesizes the poly(A) tail. Unlike in the mammalian system, yeast Pap1p is not required for cleavage of pre-mRNAs *in vitro* (Patel and Butler, 1992). Interestingly, Pap1p is phosphorylated and ubiquitinated during transition of late S to G2 phase in the cell cycle (Mizrahi and Moore, 2000). Since phosphorylation inhibits Pap1p, the poly(A) adding activity of Pap1p is cell cycle-dependent.

Fip1p physically interacts with Pap1p and is exclusively required for the polyadenylation step of 3' end processing (Preker et al., 1995). It was shown that Fip1p increases the K_M of Pap1p for RNA around 50 fold and shifts its activity to a slow and distributive mode (Zhelkovsky et al., 1998).

Pfs2p contains seven WD-repeats and deletion of a single repeat is lethal (Ohnacker et al., 2000). *pfs2* mutant strains display 3' end processing defects, indicating an essential function for Pfs2p in cleavage and polyadenylation. With a one-step affinity purification method, which exploits protein A-tagged Pfs2p, Ohnacker and co-workers showed that this protein is part of CPF. Since Pfs2p shows numerous interactions with subunits of CPF it likely promotes assembly of CPF (Ohnacker et al., 2000). Due to this finding, CPF is now commonly affinity purified in our lab via protein A-tagged Pfs2p.

Pti1p has been shown to be dispensable for both cleavage and polyadenylation (W. Hübner, unpublished). However, mutations of Pti1 or Ctk1 (a kinase that phosphorylates Ser2 of the C-terminal domain of RNAP II) affect 3' cleavage site choice and transcript abundance of particular genes (Skaar and Greenleaf, 2002). Therefore it is possible that coupling of transcription to 3' processing of pre-mRNAs is mediated by a Pti1p-containing complex (Skaar and Greenleaf, 2002). Furthermore, Pti1p is essential for yeast snoRNA 3' end maturation (Dheur et al., 2003; Nedea et al., 2003). Dheur and co-workers propose that Pti1p acts by uncoupling cleavage and

polyadenylation and functions in coordination with the Nrd1p-dependent pathway for 3' end formation of non-polyadenylated transcripts (Dheur et al., 2003).

Ref2p encodes a non-essential RNA-binding protein that has specific affinity for pyrimidine bases (Russnak et al., 1995). In addition, Ref2p was found to be essential for snoRNA 3' end termination (Dheur et al., 2003). Two-hybrid analysis and affinity purification revealed that Ref2p interacts with Glc7p (Nedea et al., 2003; Uetz et al., 2000).

Mpe1p is an RNA-binding protein and its RNA-binding property may contribute to poly(A) site selection (M. Sadowski, unpublished). By immunodepletion experiments it was demonstrated that Mpe1p is exclusively required for the polyadenylation reaction (M. Sadowski, unpublished).

Glc7p is a protein phosphatase and is the main subject of investigation of this thesis. We and others (He and Moore, 2005) found that Glc7p is required for polyadenylation of pre-mRNAs. In addition we found that Glc7p is involved in snoRNA transcription termination (Chapter 3). Furthermore, Glc7p genetically interacts with transcription elongation factors thereby affecting RNAP II during transcription elongation (Chapter 4).

Yth1p is required for both steps of pre-mRNA 3' end processing (Barabino et al., 2000). It is an RNA-binding protein and binds to pre-mRNA in the vicinity of the cleavage site, thus participating in its recognition (Barabino et al., 2000).

Ssu72p is required for 3' end cleavage of pre-mRNA but is dispensable for poly(A) addition and RNAP II termination (He et al., 2003). In addition, several other reports implicate Ssu72 in transcription termination not only of pre-mRNAs but also of snoRNAs (Dichtl et al., 2002a; Ganem et al., 2003; Nedea et al., 2003; Steinmetz and Brow, 2003). Furthermore, analyses of *ssu72-2* mutant cells revealed defects in RNAP II transcription elongation (Dichtl et al., 2002a).

Swd2p is a WD-40 repeat protein and associates with two functionally distinct multiprotein complexes: the cleavage and polyadenylation factor (CPF) and the SET1 complex (SET1C) that methylates histone 3 lysine 4 (Cheng et al., 2004; Dichtl et al., 2004). *swd2* mutant strains are defective in 3' end formation of specific mRNAs and snoRNAs. Furthermore, histone 3 lysine 4 di- and tri-methylation are adversely affected. In addition, telomeres are shortened in *swd2* mutants. However, the functions of Swd2p in CPF and SET1C are independent (Dichtl et al., 2004). In conflict to these findings, Cheng and co-workers found that cleavage and

polyadenylation of the mRNA precursor *in vitro* are normal in Swd2 depleted extracts (Cheng et al., 2004).

1.3 Transcription termination of mRNAs

Termination of pre-mRNA transcription is defined as the process by which RNAP II dissociates from the DNA template and the nascent RNA preventing further transcription. The importance of transcription termination is to avoid the inappropriate transcription of downstream genes. A failure to terminate transcription is not only a waste of energy, but also disturbs normal patterns of gene expression. In particular, transcription towards promoters of downstream genes can interfere with recruitment of RNAP II to these promoters thereby preventing the normal transcription of these downstream genes (Greger et al., 2000). In addition, the release of RNAP II molecules during termination recycles polymerases for new rounds of transcription (Dye and Proudfoot, 1999; Greger and Proudfoot, 1998).

The molecular mechanism of transcription termination by RNAP II is still not completely understood. Part of the problem is that in contrast to RNAP I and III, transcription of protein-coding genes by RNAP II terminates at ill-defined positions in the 3' UTR downstream of the poly(A) site, which directs cleavage and polyadenylation of pre-mRNAs (Proudfoot, 1989). By transcriptional run-on (TRO) analysis it was shown that transcription termination by RNAP II occurs rather randomly, at sites between 200-2000 bp downstream of the poly(A) site (Proudfoot, 1989). Nevertheless, it was demonstrated that the efficiency of transcription termination correlates with the strength of the poly(A) site (Edwalds-Gilbert et al., 1993; Orozco et al., 2002). Moreover, mutations in the poly(A) signal lead to transcription termination defects, indicating that the poly(A) signal itself is a cis-acting element required for transcription termination (Proudfoot, 1989). However, since the poly(A) signal can be far away from the actual site where RNAP II is released from the DNA template and the RNA transcript, the poly(A) site is probably not the direct trigger for the transcription termination. It remains to be elucidated whether there exists a sequence motif that triggers termination in addition to the poly(A) site.

Several reports indicate that sequence motives downstream of the poly(A) site are required for transcription termination by RNAP II (Birse et al., 1997; Connelly and Manley, 1989a; Connelly and Manley, 1989b; Yonaha and Proudfoot, 1999).

Moreover, the presence of specific G-rich sequences (pause sites) located downstream of a strong poly(A) site lead to transcriptional pausing that could facilitate termination (Yonaha and Proudfoot, 1999). This pausing of RNAP II downstream of the poly(A) site may contribute to the process of termination (Yonaha and Proudfoot, 1999). Also, studies in mammalian cells indicate that pausing might be important for transcription termination (Park et al., 2004). There is evidence that transcription termination can be divided into two steps: pausing of RNAP II and release of RNAP II from the template. Interestingly, pausing of RNAP II is CTD-independent whereas release of RNAP II is not (Park et al., 2004).

Chromatin structure across the termination region was also found to affect termination efficiency (Alen et al., 2002). Cells lacking the Chd1 chromatin remodeling factor did not only show defects in chromatin structure at termination regions but also in transcription termination itself (Alen et al., 2002). Therefore, chromatin remodeling at sites downstream of the poly(A) site might be important to facilitate pausing or even template release by RNAP II.

Besides a functional poly(A) site, there are several trans-acting factors known to be essential for transcription termination of pre-mRNAs that are involved in 3' end processing of pre-mRNAs. These include subunits of CFI A (Gross and Moore, 2001a; Minvielle-Sebastia et al., 1997), in particular Pcf11p, Rna14p and Rna15p (Birse et al., 1998b) and the CPF subunits Ssu72p (Steinmetz and Brow, 2003), Yhh1p (Dichtl et al., 2002b), Swd2p (Dichtl et al., 2004) and Ysh1p (M. Garras, personal communication). Although the mechanism, which links 3' end processing to RNA polymerase release remains unclear, the poly(A) signal is important in coupling of 3' end processing and transcription termination (see below).

1.4 Coupling of 3' end processing and transcription termination

Direct support for a connection between pre-mRNA 3' end processing and transcription termination came from studies demonstrating that the poly(A) site, a cleavage- and polyadenylation directing signal, is also a cis-element required for normal transcription termination (Connelly and Manley, 1988; Logan et al., 1987; Russo, 1995). It was shown that mutation of the poly(A) signal inhibited not only cleavage and polyadenylation at the poly(A) site but also transcription termination of RNAP II (Proudfoot, 2000b). Coupling of these two processes is further evident by

the fact that several yeast subunits of CFIA and CPF which are indispensable for cleavage and polyadenylation are also required for normal transcription termination (Aranda and Proudfoot, 2001; Proudfoot, 2004). These include the CFIA subunits Rna14p, Rna15p and Pcf11p (Birse et al., 1998b; Kim et al., 2004a; Sadowski, 2003) and CPF subunits Yhh1p (Dichtl et al., 2002b) and Ssu72p (Ganem et al., 2003; Steinmetz and Brow, 2003), Swd2p (Dichtl et al., 2004) and Ysh1p (M. Garras personal communication). In addition, the efficiencies of pre-mRNA 3' end processing and transcription termination correlate with the strength of the poly(A) site (Edwalds-Gilbert et al., 1993). Therefore, the poly(A) signal directly affects RNAP II transcription termination.

To address the mechanistic basis for the connection between termination and 3' end processing of pre-mRNAs, two models have been proposed: the allosteric model and the so-called torpedo model (Buratowski, 2005; Figure 1.3). The allosteric model suggests that upon recognition of the poly(A) site by the transcription complex, the enzymatic activity of RNAP II is modified, making it "termination-competent". Modification of RNAP II could include a conformational change or the loss of anti-termination factors, which where associated with the RNAP II. This would then lead to a loss of RNAP II processivity and eventually to transcription termination (Calvo and Manley, 2005; Greenblatt et al., 1993; Logan et al., 1987; Sadowski, 2003). The torpedo model, unlike the allosteric model, requires transcription termination to be preceded by the endonucleolytic cleavage at the poly(A) site. Cleavage of the pre-mRNA exposes an uncapped 5' end of the downstream cleavage product that serves as substrate for a 5'-to- 3' exonuclease. The exonuclease degrades the downstream cleavage product while it is still tethered to RNAP II until catching up with the RNAP II and somehow destabilizing it, ultimately terminating RNAP II (Connelly and Manley, 1988; Kim et al., 2004b; Luo and Bentley, 2004). Although both models rely on recognition of the poly(A) site, a fundamental difference is that only the torpedo model depends on cleavage of the nascent RNA to create the entry site for the 5'-to-3' exonuclease.

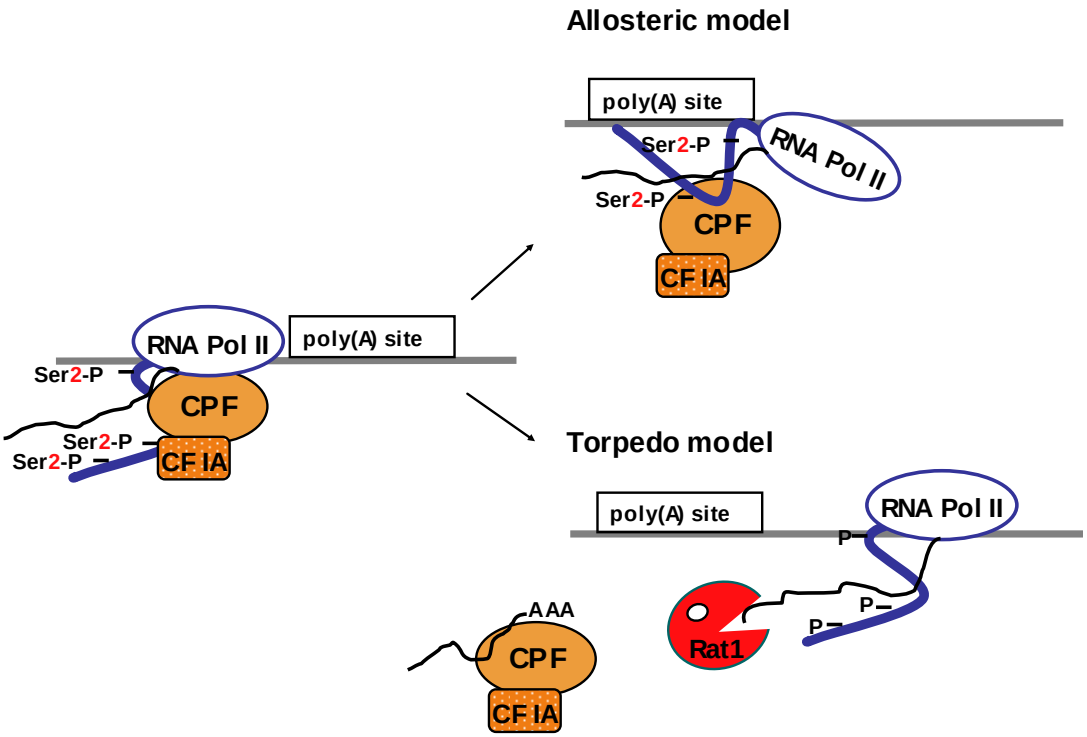


Fig. 1.3. Pre-mRNA transcription termination models

On the left hand side: the RNAP II transcription complex approaches the poly(A) site, which is required for transcription termination. The allosteric transcription termination model (right hand side, top) proposes that termination occurs cleavage-independent due to a change in the processivity of RNAP II. In the torpedo model (right hand side, bottom) a prerequisite for termination is cleavage of the pre-mRNA. The cleavage reaction generates an entry side for the exonuclease Rat1, which "torpedos" RNAP II.

There is evidence supporting either of these models. For example, the allosteric model is supported by a study showing that the 3' end processing and the termination function of Pcf11p are separable (Sadowski, 2003). Pcf11p associates directly with the nascent RNA and the CTD via its CTD-interacting domain (CID). Although the intact CID of Pcf11p is required for efficient termination, mutations in this domain have no effect on cleavage *in vitro* (Sadowski, 2003). This implies that cleavage of the pre-mRNA is not an obligatory step for transcription termination as proposed by the torpedo model. In addition, Pcf11p can force the dissociation of an *in vitro* assembled RNAP II elongation complex (Zhang et al., 2005; Zhang and

Gilmour, 2006). The authors propose that by bridging the CTD and the nascent RNA a conformational change in the CTD is induced that is transduced through Pcf11 to the nascent transcript causing transcription termination (Zhang et al., 2005; Zhang and Gilmour, 2006).

Strong evidence in support of the torpedo model comes from the recent identification of yeast Rat1 and human Xrn2, both 5'-to-3' RNA exonucleases, which are necessary for efficient transcription termination (Kim et al., 2004b; West et al., 2004). Inactivation of Rat1 does not affect cleavage or polyadenylation. However, inactivation of Rat1 results in deficient transcription termination. Moreover, the downstream cleavage products are stabilized in a rat1-1 mutant strain, suggesting that Rat1 exonuclease activity is required for the degradation of such products (Kim et al., 2004b).

A new study by Luo and co-workers (Luo et al., 2006) presents evidence that both models can be linked together. They propose a unified allosteric/torpedo model based on the following findings: (1) Rat1p degrades nascent RNA co-transcriptionally. However this degradation does not elicit termination, showing that it is not sufficient to cause RNAP II release from the template (2) Rat1p is required for the recruitment of 3' end processing factors and for correct 3' end formation. Conversely, 3' end factors are also required for normal Rat1p recruitment. Therefore it seems that Rat1p is not a dedicated termination factor, but instead, like conventional 3' end processing factors contributes to both poly(A) site cleavage and termination.

Altogether, this suggests that in addition to the exonucleolytic activity of Rat1p, Rat1p might influence termination through interactions with the cleavage and polyadenylation factors. Accordingly, Luo and co-workers proposed that Rat1 is an essential component of a RNAP II complex, which achieves cleavage at the poly(A) site, degradation of the nascent downstream RNA, and undergoes allosteric changes that promote RNAP II release from the template.

1.5 Processing and transcription termination of snoRNAs

snoRNAs are a class of non-coding RNAs that are transcribed by RNAP II. These metabolically stable RNAs, which are 60 to 300 nucleotides long, associate with a number of proteins to form small nucleolar RNPs (snoRNPs; Kiss, 2001). In the nucleolus, the majority of snoRNAs function as guide RNAs in the post-

transcriptional synthesis of 2'-O-methylated nucleotides and pseudouridines in rRNAs, tRNAs and snRNAs and most likely other cellular RNAs (Kiss, 2002). These modified nucleotides are important to ensure the proper function of rRNAs, tRNAs and snRNAs.

snoRNAs can be subdivided into two main classes according to evolutionarily conserved structural and sequence elements: the box C/D snoRNAs and the box H/ACA snoRNAs (reviewed by Kiss, 2001). While methylation of the 2'–hydroxyl groups is directed by box C/D snoRNAs, box H/ACA snoRNAs guide the conversion of uridines to pseudouridines (Kiss, 2001). They do so by base-pairing with the target RNA, specifying the nucleotides that have to be modified by the respective enzymes.

Whereas most vertebrate snoRNAs are encoded in introns of RNAP II genes, the majority of *S. cerevisiae* snoRNAs are independently transcribed monocistronic or polycistronic genes (Weinstein and Steitz, 1999). However, there are also a few intron-encoded snoRNAs in yeast (Qu et al., 1995; Villa et al., 1998).

Most precursor snoRNAs are processed at the 5' and 3' end resulting in mature, functional snoRNAs. Some snoRNAs are not processed at their 5' end, like U3, snR4 and snR13 (Samarsky and Fournier, 1999). Those, that undergo 5' end processing, as snR39b, snR40, snR47 and snR79, will loose their trimethylguanosine cap structure by cleavage by the endonuclease Rnt1p (Chanfreau et al., 1998). This event creates an entry site for Rat1 and/or Xrn1p exonucleases, which trim the 5' extensions (Lee et al., 2003). 3'end-processed snoRNAs differ in several aspects from 3' end-processed pre-mRNAs. The main difference is that mature snoRNAs are not polyadenylated. A reason for this might be the diverse genomic organization of snoRNAs that requires case-specific actions and factors. However, despite this heterogeneous organization of snoRNAs, processing of most snoRNAs relies on one general mechanism: endonucleolytic cleavage in the 5' and 3' extensions generates entry sites for exonucleases that trim the ends and release the mature snoRNAs (Allmang et al., 1999; Petfalski et al., 1998). The exonucleases that remove the 3' extensions belong to the nuclear exosome. The entry sites are often created by the Rnt1 endonuclease (Chanfreau et al., 1997; Chanfreau et al., 1998). However, Rnt1 cleavage sites are not present in all snoRNA precursors. Several snoRNAs contain sequences in their 3' portion which direct cleavage of the primary transcript by CF IA (Morlando et al., 2002). Since the cleaved precursor is not subsequently

polyadenylated, it is likely that CF IA can work in cooperation with factors different from CPF in the 3' end processing of snoRNAs.

A small number of primary snoRNA transcript is not endonucleolytically cleaved, but processed simply through trimming the ends by 3'-to-5' exonucleases (Bachellerie et al., 2002). This mode of processing requires prior transcription termination.

Seven snoRNAs in yeast are encoded in the intron of a gene. A well studied example is the U18 snoRNA, which is encoded in the intron of EFB1 (Villa et al., 1998). It was shown that two processing modes exist for these intron encoded snoRNAs. The major type of processing is splicing-dependent. The spliced host intron is first debranched and subsequently exonucleolytically trimmed resulting in the mature snoRNA. In the alternative pathway the snoRNA is endonucleolytically cut out of the intron. This second pathway is not dependent on splicing and might be the dominant form when the splicing efficiency is reduced (Villa et al., 1998). Endonucleolytic processing of snoRNAs is also required in all of the poly-snoRNA precursor transcripts that have been reported in yeast (Zagorski et al., 1988).

The transcription termination of individually transcribed snoRNAs is coupled to the recognition of cis-acting signals, as it is the case for transcription termination of pre-mRNAs. Similarly, transcription termination of snoRNAs, like in the case of termination of pre-mRNAs requires the action of CTD and RNA-interacting factors (Carroll et al., 2004; Steinmetz et al., 2001). However, the actual mechanism of snoRNA transcription termination might be different from the one of pre-mRNAs, since termination of snoRNA crucially involves the Nrd1 complex, which is not engaged in transcription termination of mRNAs. Components of the Nrd1 complex, the RNA-binding proteins Nrd1p and Nab3p, recognize specific RNA sequences in the 3' non-coding regions of snoRNAs (Carroll et al., 2004). It is to date unclear, how recognition of the termination signals is transmitted to the transcribing RNAP II and what actually forces RNAP II to terminate transcription.

1.5.1 Cis and trans-acting factors in S. cerevisiae snoRNA processing and transcription termination

1.5.1.1 Cis-acting signals

A well-characterized *cis-acting* signal in the 5' and 3' flanking regions of around 50% of snoRNAs is a stem-loop structure that is recognized by the endonuclease Rnt1p (Chanfreau et al., 1998). This hairpin structure is capped by a terminal tetraloop showing the consensus AGNN. Rnt1p recognizes and cleaves the hairpin within 13-16 bp of the tetraloop (Chanfreau et al., 2000; Wu et al., 2004). In cases where there are no Rnt1p cleavage sites, *cis-acting* signals for 3' end processing are dubious. It has been suggested that the nascent transcript is cleaved at sites that are often found downstream, close to predicted Nrd1 and Nab3 binding sites, which are required for snoRNA transcription termination (Fatica et al., 2000; Morlando et al., 2002). It is not clear whether recognition of the 3' end processing sequence is coupled to transcription termination of the snoRNA, as it is the case for pre-mRNA.

Termination of snoRNAs by RNAP II in *S. cerevisiae* requires two common RNA sequence motifs in the snoRNA 3' downstream region: GUA[AG] and UCUU (Carroll et al., 2004; Steinmetz et al., 2001; Steinmetz et al., 2006). These two motifs are bound by subunits of the Nrd1 complex, Nrd1p and Nab3p. Nrd1p binds to GUA[AG], whereas Nab3p binds to UCUU. Interestingly, Nrd1p and Nab3p binding sites are also found in the 5'UTR and 5' coding region of the Nrd1 mRNA. There, they direct the autoregulation of Nrd1 expression by premature transcription termination (Arigo et al., 2006a). As for the poly(A) signal, there is no strong consensus for these non-poly(A) terminators in *S. cerevisiae* and in addition, the arrangement of Nrd1p and Nab3p binding sites varies considerably in number and location. There are also indications that these sequence motifs work in a synergistic manner, with one motif alone being insufficient to trigger termination. Moreover, the organization of these sequences is not strongly conserved among even closely related yeasts, which indicates a high degree of genetic variability (Carroll et al., 2004). For some snoRNAs, such as snR13, snR47 and snR65, the terminator region is arranged in a bipartite fashion (Steinmetz et al., 2006). Interestingly, region I of the snR13 and snR65 terminators and the Nrd1 autoregulatory element show some sequence similarity. This suggests that regular terminators are more than just a collection of low-affinity Nrd1 and Nab3 binding sites (Steinmetz et al., 2006). Furthermore,

region II of the snR13 and snR65 terminators have sequence similarity to yeast cleavage and polyadenylation sites and can direct polyadenylation, indicating that protein-coding and non-coding RNAP II transcribed genes might use similar mechanisms to direct termination (Steinmetz et al., 2006).

1.5.1.2 Trans-acting factors

The endonuclease Rnt1p, which is the yeast homolog of RNase III, cleaves stem-loop structures in the 5' and 3' flanking regions of snRNA and many snoRNAs (Chanfreau et al., 1998). Cleavage of the stem-loops generates entry sites for exonucleases to digest 5' and 3' overhangs resulting in the mature form of the snRNA and snoRNA. The 3' overhangs are trimmed by the nuclear exosome, which consists of a number of 3' to 5' exonucleases. Since many snoRNAs do not have Rnt1p cleavage sites in their 5' and 3' flanking regions, there must be other factors that generate entry sites for these exonucleases. Indeed, snoRNAs can contain sequences in their 3' portions which direct cleavage of the primary snoRNA transcript involving CFI A (Morlando et al., 2002). However, none of the subunits of CFI A is known to have endonucleolytic activity, implying that CFI A has to work together with some other factor(s) to accomplish the cleavage reaction (Morlando et al., 2002). CFI A normally functions in combination with CPF in the cleavage and polyadenylation reaction of pre-mRNAs. It is anticipated that Ysh1p, a subunit of CPF, is the endonuclease that cleaves the pre-mRNA because it contains sequence motifs common for proteins acting on nucleic acid substrates (Callebaut et al., 2002). Morlando and co-workers showed, that a *ysh1-1* mutant strain is not involved in 3' end processing of snoRNAs (Morlando et al., 2002). However, *ysh1-1* is a very hypomorphic allele of Ysh1p that also does not show a cleavage defect in 3' end processing of pre-mRNAs (M. Garras, personal communication). We therefore consider the possibility that CFI A joins together with CPF in the 3' processing of snoRNAs.

The nuclear exosome is a protein complex consisting of at least 10 different 3' to 5' exonucleases involved in the 3' processing of rRNAs, snRNAs and many snoRNAs (van Hoof et al., 2000). It was shown that the yeast exosome mutant strains *rrp6* and *mtr4* accumulate 3'-extended polyadenylated forms of independently

transcribed snoRNAs, indicating that the nuclear exosome is required for 3' end processing of snoRNAs.

Recently, Egecioglu and co-workers (Egecioglu et al., 2006) discovered that the polyadenylation of these snoRNA processing intermediates is dependent on the activity of Trf4p and Trf5p, two variant poly(A) polymerases (Vanacova et al., 2005). Polyadenylation by Trf4p might stimulate the nuclear exosome to process the 3' ends of snoRNAs.

The 5' ends of snoRNAs are processed by the Rat1p/Rai1p nuclease complex. The complex contains 5'-to-3' exonucleases that play also a role in trimming of several ribosomal RNAs (Kim et al., 2004b). In this context, it is interesting that Rat1p is also a crucial component in transcription termination of pre-mRNAs (Kim et al., 2004b; Luo et al., 2006).

Previous studies have described a yeast pathway for transcription termination of snRNA and snoRNA genes which require the Nrd1 complex, the CTD kinase Ctk1 and the CTD of RNAP II (Conrad et al., 2000; Steinmetz et al., 2001). More recent studies have also implicated several components of APT (see below), a sub complex of CPF, in this pathway (Dheur et al., 2003; Nedea et al., 2003; Steinmetz and Brow, 2003). Mutations in any of these genes lead to read-through transcription at certain snRNA and snoRNA genes (Dheur et al., 2003; Dichtl et al., 2002a; Ganem et al., 2003; Steinmetz et al., 2001 and this study).

The RNA-binding proteins Nrd1p and Nab3p, and the helicase Sen1p have been assigned to the Nrd1 complex. Both Nrd1 and Nab3 contain a single conserved RNA recognition motif (RRM), a domain common to proteins involved in RNA processing (Conrad et al., 2000), with which they bind to the cis-acting terminator sequence elements. In addition, Nrd1p exists in two different phosphoisoforms (Conrad et al., 2000). Overexpression of Nab3p leads to the predominance of the slower migrating phosphoisoform (Conrad et al., 2000). Moreover, Nrd1p physically and genetically interacts with the CTD of RNAP II. The CTD interaction domain (CID) of Nrd1 has sequence similarity to two other yeast proteins that function at 3' ends: Pcf11p and Rtt103p (Kim et al., 2004b). They are involved in transcription termination of protein-coding genes. Therefore, it is tempting to speculate that mechanistic similarities in the process of transcription termination of snoRNAs and mRNAs may exist: Nrd1p similar to Pcf11p could dismantle the RNAP II complex from the template.

Interestingly, Nrd1p and Nab3p are also required for transcription termination of cryptic unstable transcripts (CUTs), which are RNAs that are derived from non-annotated regions of the genome (Arigo et al., 2006b; Thiebaut et al., 2006). The result of Arigo, Thiebaut and co-workers suggests that transcription termination of CUTs due to Nrd1 and Nab3 is necessary for the subsequent degradation of the CUTs by the nuclear exosome (Arigo et al., 2006a; Thiebaut et al., 2006).

It was reported previously, that the Nrd1 complex physically interacts with the nuclear exosome and stimulates the RNA degradation activity of the exosome in vitro for 3' end processing of snoRNAs (Vasiljeva and Buratowski, 2006). In the same study, additional interactions with factors of the TRAMP complex, the endonuclease Rnt1p, RNAP II subunits and the transcription elongation factor Spt5p were reported. This suggests that Nrd1 may link transcription and RNA 3' end formation of snoRNAs.

Sen1p is a Type I DNA/RNA helicase and has been implicated in processing of pre-tRNAs, pre-rRNAs and some snoRNAs (Rasmussen and Culbertson, 1998; Ursic et al., 2004; Ursic et al., 1997). Intriguingly, Sen1p is a phosphoprotein and co-purifies with Glc7p (Ho et al., 2002; Walsh et al., 2002) making it likely that Glc7p regulates the activity of Sen1p in the process of transcription termination of snoRNAs. Sen1p also interacts physically and functionally with the Rnt1p endonuclease (Ursic et al., 2004) thereby linking 3' end processing and termination.

Nrd1p, Nab3p and Sen1p are also required for maintaining normal levels of the Nrd1 mRNA. It was shown that at least Nrd1p and Nab3p do so by binding to Nrd1/Nab3 binding sites in the 5' UTR and 5' coding region of Nrd1 causing premature transcription termination (Arigo et al., 2006a).

It was proposed that the CPF complex can be subdivided into the holo-CPF complex, which includes the subunits Pta1p, Yhh1p, Ydh1p, Ysh1p, Mpe1p, Pap1p, Pfs2p, Yth1p, Fip1p and the APT sub complex comprising the subunits Syc1p, Ssu72p, Pti1p, Ref2p, Swd2p and Glc7p (Nedea et al., 2003; Figure 1.4). APT stands for associated with Pta1 to account for Pta1p's role in bridging the two sub complexes (Nedea et al., 2003). Based on the fact that Ssu72p, Ref2p and Pta1p revealed a transcription termination defect of snoRNAs, Nedea and co-workers suggested that APT is involved in the snoRNA 3' end formation (Nedea et al., 2003). This assumption was further corroborated by the findings that Swd2p and Ssu72p are required for transcription termination of snoRNAs.

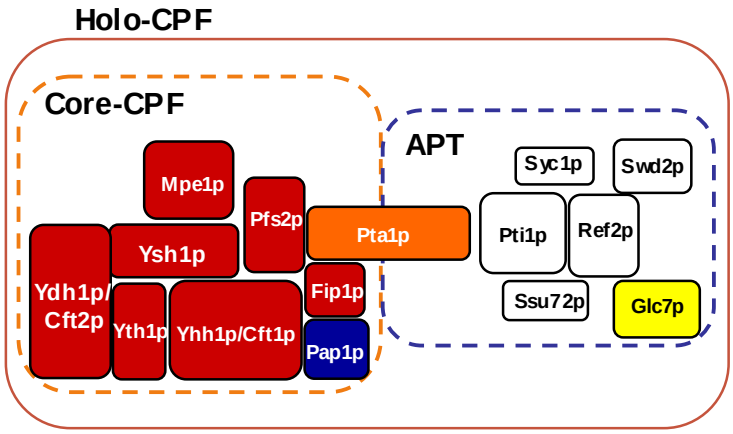


Fig. 1.4. Glc7p is part of the cleavage and polyadenylation factor subcomplex APT

Temperature sensitive *swd2* strains were defective in 3' end formation of specific snoRNAs and also some mRNAs (Dichtl et al., 2004). Interestingly, Swd2p associates with two functionally distinct multiprotein complexes: the SET1 complex, which methylates histone 3 lysine 4 and the CPF complex (Dichtl et al., 2004).

Ssu72p is a phosphatase and mediates both, termination of snoRNAs and mRNAs (Dichtl et al., 2002a; Ganem et al., 2003; Steinmetz and Brow, 2003). Steinmetz and co-workers (Steinmetz and Brow, 2003) showed that a *ssu72* mutant strain in which Nrd1-dependent termination was disrupted, also resulted in deficient poly(A)-dependent termination. Ssu72 might influence transcription termination via regulating the CTD phosphorylation, since it dephosphorylates Ser 5 of the CTD and also physically interacts with the CTD kinase Kin28p and functionally with the CTD phosphatase Fcp1p (Ganem et al., 2003; Krishnamurthy et al., 2004).

Pti1p and Ref2p were found to be required for the 3' end formation of snoRNAs, however not for 3' end processing of pre-mRNAs (Dheur et al., 2003). Additionally, Pti1p and Ref2p act as multicopy suppressors in a *pcf11-2* mutant strain (Dheur et al., 2003). Pcf11p is a subunit of CF IA and it has been reported that CF IA contributes to snoRNA and snRNA 3' end formation (Fatica et al., 2000; Morlando et al., 2002). Intriguingly, the authors suggest that Pti1p and probably also Ref2p function in the uncoupling of cleavage and polyadenylation during snoRNA 3' end formation based on the finding that overexpression of Pti1p inhibits polyadenylation.

Our own analysis on the function of Glc7p in snoRNA transcription termination extends the current knowledge on the role of the APT subunits (Chapter 3). We present evidence that Glc7p is crucial for snoRNA transcription termination.

Since no holo-CPF subunit has been shown to be involved in the 3' end formation of snoRNAs so far, the subdivision of CPF into holo-CPF and APT could not only be for physical but also for functional reasons.

Recently, the Paf1 complex, which also functions in transcription elongation and histone modification was shown to be involved in 3' end formation of snoRNAs (Sheldon et al., 2005). Deletion of Paf1p, a subunit of the Paf1 complex, leads to the accumulation of snoRNA transcripts that are extended at their 3' ends. The Paf1 complex also facilitates Nrd1 recruitment to the snR47 gene. Therefore, it was proposed that the Paf1 complex is directly involved in 3' end formation of snoRNAs (Sheldon et al., 2005).

A screen for trans-acting factors that induce readthrough of both snoRNA and mRNA terminators identified two RNAP II subunits, Rpb3p and Rpb11p (Steinmetz et al., 2006). The two subunits define a discrete surface at the trailing end of RNAP II (Steinmetz et al., 2006). This suggests that the signal for RNAP II to release the template and thus to terminate transcription is mediated by the Rpb3/Rpb11 heterodimer. In addition, this finding supports the idea that protein-coding and non-coding RNAP II-transcribed genes might use a similar mechanism to direct transcription termination. The fact that some factors involved in cleavage and polyadenylation of pre-mRNAs are required for 3' end formation of snoRNA further substantiates that there are parallels between the poly(A) and the non-poly(A) termination mechanisms.

1.6 The C-terminal domain of RNAP II, an important player in coupling transcription to capping, splicing and 3' end processing of pre-mRNAs

A striking feature that distinguishes RNAP II from RNAP I and III, is the extended carboxyl-terminal domain (CTD) of the largest RNAP II subunit Rpb1. The CTD consists of multiple heptapeptide repeats of the consensus sequence $Y_1S_2P_3T_4S_5P_6S_7$ (Prelich, 2002). The number of repeats increases with the complexity of the organism, for example, the yeast CTD has 26 repeats, whereas the human CTD

has 52 repeats (Prelich, 2002). In addition, the number of repeats that exactly match the consensus sequence varies among species (Prelich, 2002). The inherently unstructured CTD protrudes from the RNAP II enzyme and therefore can be easily accessed by numerous factors that bind to it (Prelich, 2002). As will be discussed below, the factors that bind the CTD are mostly RNA processing factors and elongation factors that act co-transcriptionally to mature the RNA and allow the transcription of the respective gene. Therefore, the CTD is crucial in coupling RNA processing events to the transcription of pre-mRNAs by attracting and transporting many factors that are involved in the production of mature mRNAs.

The CTD can be modified by phosphorylation, glycosylation and cis/trans isomerization of prolines (Meinhart et al., 2005). Modification of the CTD markedly affects its conformation and ability to associate with factors that are involved in transcription elongation, RNA processing and transcription termination. Therefore, modification of the CTD is important for the coordination of transcription events, and different modification states of the CTD are characteristic of different transcriptional stages (Prelich, 2002).

Phosphorylation of the CTD occurs predominantly at serine 2 (Ser2) and serine 5 (Ser5) (Dahmus, 1996). There are basically three phosphorylation states of the CTD: the unphosphorylated CTD, the CTD phosphorylated at Ser2 and the CTD phosphorylated at Ser5. The current paradigm is that different phosphorylation patterns predominate at different stages in the transcription cycle and that different proteins bind to specific phosphorylated forms of the CTD (Figure 1.5). This model is based on results from chromatin immunoprecipitation (ChIP) experiments, which suggested that the CTD is unphosphorylated during the assembly of the transcription complex at the promoter. Following transcription initiation, the level of Ser5 phosphorylation increases rapidly. It remains unclear whether the levels of Ser5 phosphorylation during transcription elongation remain constant or decrease towards the 3' end of the gene. Ser2 phosphorylation levels appear to increase during elongation, reaching a peak near the poly(A) site after which the level of Ser2 phosphorylation drops again (Ahn et al., 2004; Kim et al., 2004a; Komarnitsky et al., 2000; Morris et al., 2005).

Probably at least some repeats remain unphosphorylated. In addition, Ser5 phosphorylation levels do not drop to zero during elongation indicating that not all repeats of the CTD are phosphorylated in the same way (Ahn et al., 2004).

Furthermore, different findings indicate that during the elongation stage repeats, which are phosphorylated at both Ser2 and Ser5 must exist. First, Ctk1p, the CTD Ser2 kinase, efficiently uses Ser5 phosphorylated repeats to generate Ser2/Ser5-phosphorylated repeats *in vitro* (Jones et al., 2004). Second, the histone methyltransferase Set2p, which remodels chromatin during transcription elongation, requires repeats phosphorylated on both Ser2 and Ser5 for optimal binding (Kizer et al., 2005; Li et al., 2005). Finally, the monoclonal Antibody (mAb) H5, which is used to detect Ser2 phosphorylation, binds better to repeats concomitantly phosphorylated at Ser2 and Ser5 (Jones et al., 2004). This indicates that what was interpreted as Ser2 phosphorylation might actually be phosphorylation on both Ser2 and Ser5.

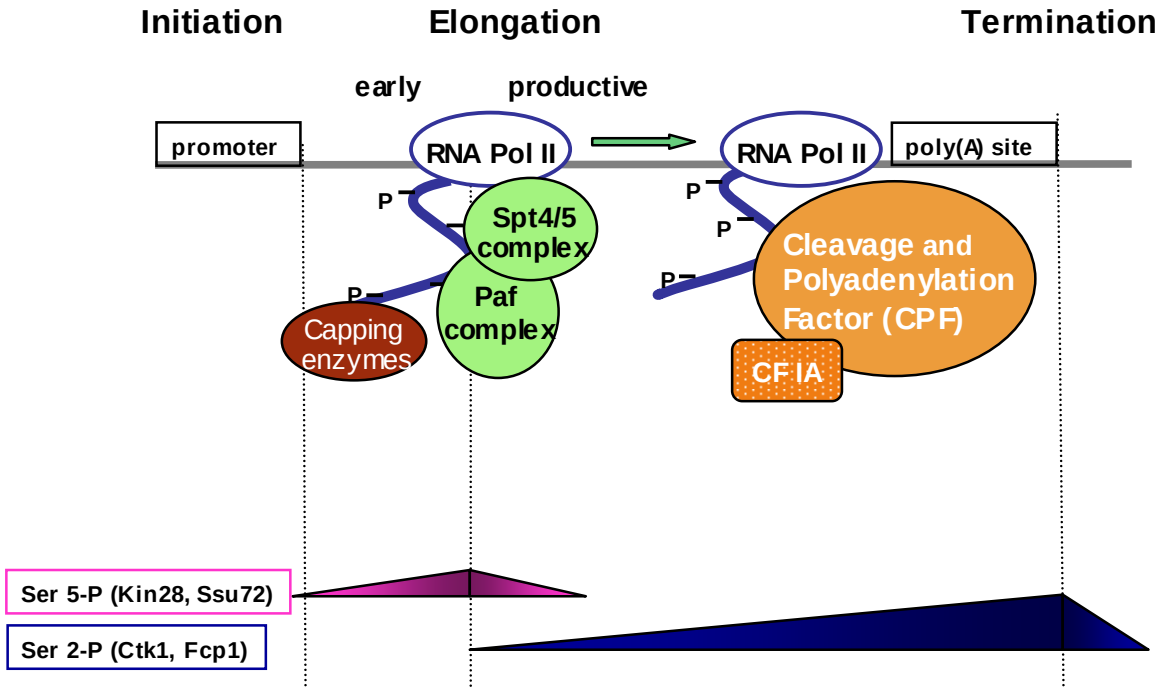


Fig. 1.5. Linking RNAP II transcription and RNA processing: a central role for the CTD
The phosphorylation pattern of the CTD changes during different transcription phases indicated by the purple (for Ser 5 phosphorylation) and blue (for Ser 2 phosphorylation) bars. As a consequence distinct processing factors (capping enzymes, the Spt4/Spt5 complex, the Paf complex, CPF and CFIA) are attracted to the CTD during the different phases of transcription.

Several CTD kinases and phosphatases are know to modify the phosphorylation state of the CTD of RNAP II (Phatnani and Greenleaf, 2006). The

known CTD phosphatases include Ssu72p, small CTD phosphatases (SCPs) and Fcp1p. Kinases required for CTD phosphorylation are Kin 28 and Ctk1.

Ssu72 has been shown to be a Ser5 phosphatase (Krishnamurthy et al., 2004). The depletion of Ssu72 and concomitant elevation of Ser5 phospho-levels inhibit transcription *in vitro* (Krishnamurthy et al., 2004). Interestingly, Ssu72p is a member of the CPF complex and has functions in all stages of transcription from initiation over elongation to termination (Dichtl et al., 2002a; Ganem et al., 2003; Pappas and Hampsey, 2000; Steinmetz and Brow, 2003), (Meinhart et al., 2005). The essential role of Ssu72p in 3' end processing of pre-mRNAs however, is independent of its phosphatase activity (Krishnamurthy et al., 2004).

Small CTD phosphatases (SCPs), which are specific to higher eukaryotes, preferentially dephosphorylate Ser5 (Yeo et al., 2003).

FCP1 can dephosphorylate both Ser2 and Ser5 *in vitro* (Lin et al., 2002). However, there are discrepancies as to which substrate is preferred *in vivo*. Mutations in yeast Fcp1p preferentially cause an increase in the levels of phospho-Ser2 (Cho et al., 2001). Dephosphorylation of the CTD by Fcp1p might facilitate recycling of the hyperphosphorylated form of the polymerase for a new round of transcription (Cho et al., 1999). FCP1 might also be required to regulate phosphorylation of the CTD during transcription (Cho et al., 2001). CTD dephosphorylation by Fcp1 can also be stimulated by Ess1, a peptidylprolyl isomerase (Kops et al., 2002; Xu et al., 2003). Ess1 catalyzes the conversion of the Ser-Thr-Pro moiety from the *cis* to the *trans* conformation (Kops et al., 2002).

In *S. cerevisiae*, four cyclin-dependent kinases have been identified as being important for transcription, however only two of them have unambiguously been shown to phosphorylate the CTD. Among the latter is the Kin28 (Cdk7) subunit of TFIIH, which phosphorylates the CTD on Ser5, when transcription is initiated (Rodriguez et al., 2000). Ctk1 (Cdk9) associates with elongating RNAP II and phosphorylates the CTD on Ser2 (Cho et al., 2001).

Capping of pre-mRNAs occurs soon after transcription initiation and before other processing events. Phosphorylation of Ser5 of the CTD heptapeptide repeats is important for the recruitment and the enzymatic activity of the mRNA capping enzyme guanylyltransferase (Cho et al., 1998; Cho et al., 1997; Ho et al., 1998; Komarnitsky et al., 2000; Rodriguez et al., 2000). Interestingly, levels of the polyadenylation factor Pta1p are reduced in *kin28* mutants in which Ser5

phosphorylation is poor, indicating that this polyadenylation factor might be recruited to the transcription complex already at early stages of transcription. Several other polyadenylation factors remain stable in the *kin28* mutant (Rodriguez et al., 2000). In addition, Pta1p binds specifically to the phosphorylated CTD. This suggests that polyadenylation and transcription are coupled through this interaction (Rodriguez et al., 2000).

Phosphorylation of Ser2 by Ctk1 is required for the recruitment of polyadenylation factors (Rna14p, Rna15p, Yhh1p and Ydh1p) to the 3' end of genes *in vivo* (Ahn et al., 2004). Also Pcf11p, a subunit of cleavage factor CF IA, was found to bind efficiently to a Ser2-phosphorylated CTD peptide *in vitro* (Licatalosi et al., 2002). Furthermore, deletion of Ctk1 leads to changes in polyadenylation site choice, presumably due to the loss of co-transcriptional polyadenylation, but does not affect transcription termination *in vivo* (Ahn et al., 2004). Ahn and co-workers also showed that Ctk1 and therefore Ser2 phosphorylation is not required for the recruitment of the elongation factors Spt5p, Spt4p, Spt6p and the PAF complex (Ahn et al., 2004). In contrast, the chromatin remodeling factors Set1p and Set2p bind to the phosphorylated CTD (Krogan et al., 2003a; Krogan et al., 2003b; Ng et al., 2003b). As mentioned above, the optimal binding conditions, at least for Set2p, might be heptapeptide repeats phosphorylated at both serines (Kizer et al., 2005; Li et al., 2005).

Interestingly, deletion of Ctk1 leads to a defect in snoRNA transcription termination (Steinmetz et al., 2001), suggesting that phosphorylation of Ser2 is necessary for efficient snoRNA transcription termination.

The RNAP II CTD itself is essential for 3' end cleavage of metazoan pre-mRNAs (McNeil et al., 1998), since deletion of the yeast CTD reduces the efficiency of cleavage at poly(A) sites and the length of poly(A) tails. This suggests that it helps to couple 3' end formation with transcription (Licatalosi et al., 2002).

Normal mRNA splicing and 3' end processing are disrupted in mammalian cells expressing RNAP II lacking the CTD (Fong and Bentley, 2001; McCracken et al., 1997). In yeast, splicing appears unaffected when the CTD is deleted form RNAP II (Licatalosi et al., 2002). Moreover, it was shown that whereas the carboxyl-terminal part of the mammalian CTD consisting of heptads 27-52, supports capping, splicing and 3' end processing, the amino-terminus supports only capping (Fong and Bentley, 2001). This indicates that the CTD as a whole stimulates all three major pre-mRNA

processing steps but that different regions of the CTD serve distinct functions in the different subprocesses.

Although many functional aspects concerning the CTD of RNAP II have been solved, still many questions remain open, e.g.: what initiates CTD phosphorylation? What causes the dissociation of factors from the CTD? Is there a particular order of factors bound along the length of the CTD?
Therefore, the CTD of RNAP II continues to be an exciting field of investigation.

1.7 Transcription elongation

Transcription of a gene by RNAP II can be divided into three main phases: initiation, elongation and termination of transcription (Saunders et al., 2006). For many years, the focus of how gene expression is regulated, was concentrated on the process of transcription initiation. The subsequent transcription elongation was thought of as the simple addition of ribonucleotide triphosphates to the growing RNA chain (Saunders et al., 2006). Only since recently, it is appreciated that also during transcription elongation mechanisms exist that allow the regulation of gene expression (Saunders et al., 2006; Sims et al., 2004b). This is due to the identification of a number of transcription elongation factors that affect mRNA production at particular stages during transcription (Saunders et al., 2006).

Transcription elongation is divided into three distinct stages: promoter escape (or promoter clearance), promoter-proximal pausing and productive elongation (Saunders et al., 2006). Each stage is defined by a marked difference in the stability and behavior of the RNAP II transcription complex as well as by a distinct repertoire of factors that associate with it (Saunders et al., 2006).

During promoter escape, some of the general transcription factors dissociate from the RNAP II transcription complex and RNAP II breaks its contacts to promoter-sequence elements (Saunders et al., 2006; Sims et al., 2004a). These actions lead to the release of the RNAP II transcription complex from the promoter and the start of transcription of the gene. However, to acquire full processivity, the RNAP II transcription complex has yet to be further transformed. Full transformation into a highly productive transcription elongation complex is only achieved after a process known as promoter-proximal pausing (Saunders et al., 2006; Sims et al., 2004b). Promoter-proximal pausing is a phenomenon where RNAP II pauses at the 5’ region

of the transcription unit and only progresses efficiently into productive elongation upon stimulation by appropriate signals. Pausing can be considered as a checkpoint before committing to productive elongation. During this pausing, capping enzymes are recruited to the stalled transcription elongation complex and the nascent RNA is capped (Saunders et al., 2006; Sims et al., 2004a). Escape from the pause site and promoter escape are the rate-limiting steps in RNAP II transcription elongation (Lis, 1998). In addition, gene expression can be regulated at these points (Lis, 1998).

RNAP II encounters only a few obstacles during the productive elongation stage. However, to smoothly operate during this stage, RNAP II is dependent on transcription elongation factors that help to overcome intrinsic arrest sites or remodel the chromatin structure that would otherwise impair further transcription (Saunders et al., 2006; Sims et al., 2004a).

A crucial player in all stages of transcription elongation is the CTD of RNAP II (Hirose and Manley, 2000; Orphanides and Reinberg, 2002; Proudfoot et al., 2002). It acts as a platform that attracts and binds many factors involved in transcription elongation as well as factors involved in mRNA maturation, surveillance and export (Hirose and Manley, 2000; Orphanides and Reinberg, 2002; Proudfoot et al., 2002) (see also Chapter CTD). The highly dynamic phosphorylation pattern of the CTD acts as a code that ensures the stage-specific recruitment of a subset of transcription elongation factors (Buratowski, 2003; Conaway et al., 2000; Howe, 2002). Once recruited, they travel along with RNAP II and act co-transcriptionally to modulate the processivity of RNAP II or remodel chromatin until they are no longer needed (Maniatis and Reed, 2002; Orphanides and Reinberg, 2002).

1.7.1 Elongation factors that modulate the catalytic activity of RNAP II

Many elongation factors that modulate the catalytic activity of RNAP II do so by counteracting or alleviating three major impediments to transcription elongation: transcriptional pausing, transcriptional arrest and transcription termination (Sims et al., 2004a). Transcriptional pausing occurs when the RNAP II halts the addition of NTPs to the nascent RNA transcript for a time before resuming productive elongation on its own (Sims et al., 2004a). Transcriptional arrest can be defined as an irreversible halt to RNA synthesis, whereby the RNAP II cannot resume productive elongation without accessory factors (Sims et al., 2004a). Transcription termination is described

as the process in which RNAP II is released from the template DNA and the RNA transcript thereby ending the elongation stage of transcription.

1.7.1.1 Factors that influence transcriptional pausing

Factors that influence transcriptional pausing of RNAP II are the general transcription factor TFIIF, the positive transcription elongation factor b (P-Tefb) and the DRB-sensitivity-inducing factor (DSIF).

1.7.1.1.1 TFIIF

Of all the general transcription factors, only TFIIF can be found in the RNAP II transcription elongation complex (Zawel et al., 1995). It was shown that the human general transcription factor TFIIF diminishes the time RNAP II is paused and stimulates the rate of RNAP II transcription elongation (Tan et al., 1994). However, TFIIF does not move along with RNAP II, but re-associates with the polymerase when encountering a block to elongation. It was postulated that the transient association of TFIIF with paused RNAP II induces a conformational change in the polymerase necessary for optimal elongation. Once this is accomplished, TFIIF is released from the transcription elongation complex (Zawel et al., 1995). Moreover, TFIIF directly interacts with and stimulates the CTD Ser2 phosphatase Fcp1 (Kamada et al., 2003). Fcp1 stimulates transcription elongation *in vitro* (Mandal et al., 2002). The stimulatory activity of TFIIF and Fcp1p on elongation is additive (Mandal et al., 2002). Interestingly, TFIIF also associates with the transcription elongation factors Spt5 and components of the PAF complex (Lindstrom et al., 2003; Shi et al., 1997).

1.7.1.1.2 P-TEFb

Treatment of mammalian cells with the nucleotide analog DRB (5,6-dichloro-1-b-D-ribofuranosylbenzimidaole) results in the formation of truncated transcripts indicating that RNAP II ceased transcription before the end of the gene (Chodosh et al., 1989). P-TEFb (positive transcription elongation factor b) was originally identified as a factor that was able to alleviate the effect of DRB on transcription of long transcripts *in vitro* (Marshall et al., 1996). It consist of the Cdk9 kinase that

associates with different cyclins (Peng et al., 1998). The yeast homolog of Cdk9 is the kinase Ctk1. P-TEFb phosphorylates Spt5 and Ser2 of the RNAP II CTD and this phosphatase activity is important for the elongation activity of P-TEFb (Fujinaga et al., 2004; Yamada et al., 2006). Spt5 together with Spt4 forms the DSIF transcription elongation complex (Hartzog et al., 1998; Wada et al., 1998a). DSIF contributes to promoter-proximal pausing of RNAP II. P-TEFb counteracts the negative effect of DSIF on transcription elongation by phosphorylating Spt5 and full-length transcription is resumed (Yamada et al., 2006). Phosphorylation of Ser2 of the CTD by p-TEFb is a further hallmark of the transcription elongation phase of transcription. In addition, inhibition of P-TEFb culminates in incorrect RNA 3' end processing, consistent with the important role for the Ser2-phosphorylated CTD in the recruitment of 3'end processing factors (Ahn et al., 2004; Ni et al., 2004).

1.7.1.1.3 DSIF

DRB-sensitivity-inducing factor (DSIF) was identified based on its ability to confer DRB sensitivity to a transcription reaction (Wada et al., 1998a). The heterodimeric DSIF complex is composed of the proteins Spt5 and Spt4. DSIF was originally reported to inhibit transcription elongation, more recent findings however also implicate a positive function for Spt4 and Spt5 in transcription elongation (Lindstrom et al., 2003; Rondon et al., 2003). *In vitro* transcription assays and genetic studies in yeast revealed Spt4 as positive transcription elongation factor (Rondon et al., 2003). Furthermore, Spt4 can antagonize pausing of RNAP II imposed by the chromatin-remodeling factor Isw1p (Morillon et al., 2003a).

Studies on Spt5 showed that early in transcription Spt5 is inhibitory to transcription, whereas later in transcription, after promoter-proximal pausing, Spt5 stimulates the transcription process by preventing premature termination and pausing (Bourgeois et al., 2002; Wada et al., 1998b). Biochemical and genetic evidence supports interactions between Spt4, Spt5 and RNAP II (Wada et al., 1998a). Further interaction studies demonstrated that DSIF genetically and physically associates with TFIIF as well as with the chromatin-related factors Spt6, FACT, Chd1 and the PAF complex (Costa and Arndt, 2000; Hartzog et al., 1998; Krogan et al., 2002; Lindstrom et al., 2003; Mueller and Jaehning, 2002; Squazzo et al., 2002).

1.7.1.2 Factors that influence transcriptional arrest

Arrested RNAP II complexes resume productive elongation via an evolutionarily conserved mechanism that requires cleavage of the RNA transcript (Sims et al., 2004a). TFIIS was identified as factor that promotes RNAP II read-through at arrest sites (Conaway et al., 2003) and therefore stimulates elongation by reducing RNAP II pausing. In addition, it enhances the cleavage reaction, which is an intrinsic property of RNAP II itself (Orlova et al., 1995). Recently, the structure of RNAP II in combination with TFIIS could be solved (Kettenberger et al., 2003). From this structure, Kettenberger and colleagues could determine that the active site of RNAP II undergoes extensive structural changes during TFIIS binding and these structural changes are consistent with a realignment of the RNA in the active center (Kettenberger et al., 2003). The yeast homolog of TFIIS is termed Ppr2p. *ppr2* mutants were shown to be sensitive to the drug 6-azauracil (6AU). Sensitivity to 6AU is a hallmark of transcription elongation mutants. In addition, TFIIS genetically interacts with elongation factors like DSIF and also with the PAF complex, FACT, Spt6 and SWI/SNF, which are chromatin remodelers (see below; Costa and Arndt, 2000; Davie and Kane, 2000; Hartzog et al., 2002; Lindstrom and Hartzog, 2001; Orphanides et al., 1999).

1.7.2 Elongation factors that remodel chromatin

In eukaryotes, DNA is packaged into chromatin. An important structural unit of chromatin is the nucleosome. The nucleosome consists of an octameric complex of histone proteins formed of an H3-H4 tetramer and two H2A-H2B dimers around which 147 bp of DNA are wrapped (Saunders et al., 2006). A linker DNA connects the nucleosomes. H1, a fifth histone protein can bind to the linker DNA and helps to stabilize higher order chromatin (Khorasanizadeh, 2004).

Chromatin represents a major block for RNAP II transcription (Sims et al., 2004a) and therefore has to be remodeled during transcription elongation. Models that explain how RNAP II can transcribe in a chromatin environment have been put forward. They propose that both nucleosome mobilization and histone depletion occur as RNAP II progresses along the chromatin template (Studitsky et al., 2004). However, it still remains to be elucidated how the elongation factors that remodel or modify the chromatin structure work together and how histone modifications

influence the chromatin structure and thus transcription elongation (Sims et al., 2004a). Elongation factors that remodel chromatin can be further subdivided in ATP-dependent remodeling factors, histone chaperones and factors that modify histones.

1.7.2.1 ATP-dependent remodeling

SWI/SNF and Chd1 are two factors that share a conserved ATPase domain (Lusser and Kadonaga, 2003). SWI/SNF is an ATP-dependent chromatin-remodeling complex whose ATPase activity is required for transcription on nucleosomal templates (Narlikar et al., 2002; Sullivan et al., 2001). The ATPase Chd1 remodels nucleosomes *in vitro* and appears to function in both elongation and termination (Alen et al., 2002; Tran et al., 2000). Chd1 interacts genetically with the SWI/SNF complex (Tran et al., 2000) and physically with the PAF complex, DSIF and FACT (Kelley et al., 1999; Krogan et al., 2002; Simic et al., 2003). All these factors have been implicated in transcription elongation (see below).

1.7.2.2 Histone Chaperones

Two factors in particular have been linked with histone disassembly and reassembly: FACT and Spt6. FACT facilitates the displacement of H2A-H2B dimers during *in vitro* transcription and facilitates RNAP II elongation through chromatin (Belotserkovskaya et al., 2003; Sims et al., 2004a). Spt6 interacts physically with histones *in vitro*, most strongly with H3 and H4, and also helps to disassembly the H2A-H2B dimer (Bortvin and Winston, 1996). After RNAP II has passed by a nucleosome, FACT and Spt6 help to reassemble the nucleosome (Saunders et al., 2006). Reassembly of the nucleosome after RNAP II transcription is important for preventing aberrant transcription initiation from cryptic promoter sites that contain a TATA element and a proximal initiation site (Saunders et al., 2006).

1.7.2.3 Histone modification factors

Histone proteins are posttranslationally covalently modified by a wide variety of enzymes. The highly accessible N-terminal histone tails undergo several covalent modifications including acetylation, methylation, ubiquitination, phosphorylation and

ADP-ribosylation (Vaquero et al., 2003; Zhang, 2003). Histones of genes that are transcriptionally active are highly acetylated in comparison to histones of genes that are silent. Histone acetylation destabilizes chromatin structure by disrupting internucleosome associations as well as histone tail interactions with linker DNA (Sims et al., 2004a). Therefore, acetylation of histones is important for productive transcription (Saunders et al., 2006). Histone lysine methylation at specific lysines appears to play a role in establishing both short-and long-term transcriptional regulation (Sims et al., 2004a). Silent or repressed regions co-map with methylation at histone lysine residues H3-K9, H3-K27 and H4-K20 (Sims et al., 2004a). In contrast, transcriptionally active domains are typically associated with methylation at H3-K4, H3-K36 and H3-K79 (Lachner et al., 2003; Sims et al., 2003). These lysines are methylated by specific methyl transferases called Set1p and Set2p.

Set1p is a specific histone H3-H4 methyltransferase in yeast. It associates with RNAP II CTD and the PAF complex (Krogan et al., 2003a; Ng et al., 2003a) and most interestingly also with the CPF complex (Dichtl et al., 2004). Set1-mediated histone H3-H4 methylation occurs at promoters and within the coding region of active genes in yeast (Ng et al., 2003b).

Yeast Set2p is an H3-K36-specific histone methyltransferase. Set2 associates with the hyperphosphorylated form of RNAP II CTD (Krogan et al., 2003b). It was further demonstrated that association of Set2 and H3-K36 methylation within coding regions is dependent on the PAF complex (Krogan et al., 2003b).

The PAF complex consist of five subunits, Paf1, Ctr9, Cdc73, Rtf1 and Leo1 and is conserved from yeast to humans (Krogan et al., 2002; Mueller and Jaehning, 2002; Squazzo et al., 2002). Paf1 was originally identified as a protein associated with RNAP II (Shi et al., 1996). The PAF complex has a crucial role in the regulation of histone monoubiquitylation and methylation of several histones. The PAF complex, in particular its Rtf1 component, is required for the recruitment of the H2B monoubiquitylation enzyme and therefore for monoubiquitylation of H2B (Wood et al., 2003). Histone H2B monoubiquitylation was shown to be a prerequisite for H3-K4 and H3-K79 methylation (Ng et al., 2003a; Wood et al., 2003). Similarly, the PAF complex is required for recruitment of Set1 and Set2 in a manner dependent upon the phosphorylation state of the CTD of RNAP II and also for correct H3-K4, H3-K36 and H3-K79 methylation (Krogan et al., 2003a; Krogan et al., 2003b; Ng et al., 2003b; Wood et al., 2003). Apart from their many functions in histone modification,

the PAF complex subunits revealed a wide range of transcription elongation phenotypes, including sensitivity to 6AU. In addition, the PAF complex associates with transcription elongation factors, indicating that it is involved in transcription elongation, too (Costa and Arndt, 2000; Mueller and Jaehning, 2002; Rondon et al., 2003; Squazzo et al., 2002). However, a recent study also found functions for the PAF complex which are independent of actively transcribing RNAP II (Mueller et al., 2004). In addition, loss of the Paf1 factors causes a reduction of RNAP II CTD Ser2 phosphorylation and also shortened poly(A) tails, suggesting that the complex facilitates linkage of transcriptional and posttranscriptional events (Mueller et al., 2004). Yet an other function for the PAF complex has been reported recently. It was shown that the PAF complex is required for the 3' end formation of snoRNAs, thus probably linking transcription elongation of snoRNAs to their 3' end formation (Sheldon et al., 2005).

1.8 Protein phosphatases

Reversible protein phosphorylation is involved in the regulation of cellular processes as diverse as glycogen metabolism, muscle contraction, ion channel regulation, transcription, translation and the cell cycle (Depaoli-Roach et al., 1994). Protein phosphatases are signal-transducing enzymes that dephosphorylate cellular phospho-proteins (Barford, 1995). The dephosphorylation of serine, threonine and tyrosine residues within proteins is catalyzed by three families of protein phosphatases including two families of Ser/Thr phosphatases and one protein tyrosine phosphatase family (Barford, 1996). Diversity of structure within a family is determined by targeting and regulatory subunits and domains. Structural studies of these enzymes have revealed that although the two families of protein Ser/Thr phosphatases are unrelated in sequence, the architecture of their catalytic domains is remarkably similar and distinct from the protein tyrosine phosphatases (Barford, 1996).

Surprisingly, protein phosphatases are outnumbered by a large amount of protein kinases in the cell. Over the past decade an understanding has emerged of how such a small number of phosphatases is able to dephosphorylate thousands of proteins while allowing the level of phosphorylation of each of these proteins to be regulated independently. Many of these questions have been answered by studies of one of the

major eukaryotic protein phosphatases: protein phosphatase 1 (PP1). The PP1 catalytic subunit can form complexes with > 50 regulatory subunits in a mutually exclusive manner. The formation of these complexes converts PP1 into many different forms, which have distinct substrate specificities, restricted subcellular locations and diverse regulation. This allows numerous cellular functions that rely on PP1 to be controlled by independent mechanisms (Cohen, 2002).

The S. cerevisiae genome encodes ~32 protein phosphatases (Sakumoto et al., 1999). Three of the 32 phosphatase are essential for cell viability (Sakumoto et al., 1999). One of them is the serine/threonine-specific phosphatase Glc7p (Feng et al., 1991).

1.8.1 Glc7p

Glc7p is the only yeast homolog of the mammalian PP1 (Feng et al., 1991) and regulates both nuclear and cytoplasmic processes. The Glc7 protein has a sequence that is 81% identical with the rabbit protein phosphatase 1 catalytic subunit (Feng et al., 1991). The molecular weight of Glc7p is 36 kDa. It belongs to the few genes in *S. cerevisiae* that have an intron. Exon 1 of Glc7p spans nucleotides 1 – 177 and Exon 2 nucleotides 703 – 1464 (Figure 1.6 a). Glc7p has been implicated in processes such as glycogen accumulation, carbohydrate metabolism, translation initiation (Ashe et al., 2000), transcription, mitosis, meiosis, vesicular trafficking, cell polarity, histone modification and ion homeostasis (Stark, 1996). Time-lapse fluorescence microscopy revealed that GFP-Glc7p localizes predominantly in the nucleus throughout the mitotic cell cycle, with the highest concentrations in the nucleolus (Bloecher and Tatchell, 2000). GFP-Glc7 was also observed in a ring at the bud neck, which was dependent on functional septins (Bloecher and Tatchell, 2000).

Originally, Glc7p has been identified as a gene that complements a defect in glycogen accumulation, hence its name: Glc7 for Glycogen-deficient (Feng et al., 1991). Later, it was found that Glc7p regulates the phosphorylation of Gsy2, the major isoform of glycogen synthase (Anderson and Tatchell, 2001). Gac1p is the regulatory and targeting subunit of Glc7p in the dephosphorylation of Gsy2p (Wu et al., 2001).

Glc7p is involved in several different steps during mitosis. Distinct mutations in Glc7p lead to an arrest in the cell cycle (MacKelvie et al., 1995). It was proposed

that Sds22p is the regulatory subunit of Glc7p in this process and that Sds22p functions positively with Glc7p to promote dephosphorylation of nuclear substrates required for faithful transmission of chromosomes during mitosis (Peggie et al., 2002).

Moreover, Glc7p regulates the mitotic phosphorylation of histone H3 at serine 10 in *S. cerevisiae* (Hsu et al., 2000). The phosphatase activity of Glc7p is opposed by the kinase Ipl1p, which phosphorylates H3 S10 (Hsu et al., 2000). It is believed that phosphorylation of histone H3 is required for chromosome segregation (Hsu et al., 2000). A function for Glc7p in meiosis was proposed, based on the fact that Glc7p counteracts Mek1p kinase activity, which is required for the pachytene checkpoint, and checkpoint-induced pachytene arrest (Bailis and Roeder, 2000).

Glc7p also has functions in glucose repression in S. cerevisiae (Tu and Carlson, 1994). Glucose repression is a signaling pathway that suppresses transcription of genes required under glucose starvation conditions. Reg1p was identified as the regulatory subunit of Glc7p in glucose repression (Tu and Carlson, 1995). The Reg1p-Glc7p complex antagonizes the action of the protein kinase Snf1p in the transcription of genes required for non-carbon source utilization (Young et al., 2002). Activation of Snf1p, when glucose is limiting, requires phosphorylation of a threonine residue at the C-terminus of Snf1p. It was shown that Glc7p regulates the phosphorylation state of Snf1p .

Reg1p interacts with the 14-3-3 proteins Bmh1 and Bmh2. Together they regulate the glucose-induced degradation of maltose permease (Mayordomo et al., 2003) and also have roles in maintaining glucose repression (Dombek et al., 2004). Reg1p-Glc7p is required for the glucose-induced degradation of fructose-1,6-bisphosphatase in the vacuole (Cui et al., 2004).

Interestingly, a defect in processing of rRNA was also reported for a conditional Glc7p deletion strain (Peng et al., 2003). Deletion of Glc7p results in the accumulation of 35S pre-RNA, indicating a defect in A2 cleavage of the pre-rRNA (Peng et al., 2003). The delay in A2 cleavage may be a consequence of either altered snoRNA processing and/or a reduction in the abundance of ribosomal proteins (Peng et al., 2003).

Two regulatory proteins of Glc7p have been described (Garcia-Gimeno et al., 2003; Nigavekar et al., 2002). One of them, Glc8p, is an activator of Glc7p (Nigavekar et al., 2002). However, Glc8p induces the activity of Glc7p only during

stationary phase, because the expression of Glc8p is repressed by glucose (Nigavekar et al., 2002). The other regulatory protein, Ypi1p, is an inhibitor of the phosphatase activity of Glc7p (Garcia-Gimeno et al., 2003).

1.8.2 Description of temperature-sensitive glc7 mutant alleles used for this thesis

During the course of this thesis we worked with four temperature-sensitive (ts) glc7 mutant alleles: *glc7-10*, *glc7-12*, *glc7-13* and *glc7-5* and the corresponding isogenic W303 WT strain (Figure 1.6 b). These strains were all generated in the laboratory of Prof. Mike Stark, University of Dundee. The isogenic WT background is MATa his3 leu2 uro3 ade2 glc7::LEU2 trp1::GLC7::TRP1. The reported phenotypes of these glc7 mutant alleles will be discussed in the following paragraphs.

1.8.2.1 glc7-10

The temperature-sensitive *glc7-10* allele displays aberrant bud morphology and an abnormal actin cytoskeleton at the restrictive temperature (Andrews and Stark, 2000). At 37ºC, *glc7-10* strains accumulate a high proportion of budded cells with an unmigrated nucleus, duplicated spindle pole bodies, a short spindle, delocalized cortical actin and 2C DNA content indicating a cell cycle block prior to the metaphase to anaphase transition (Andrews and Stark, 2000). Pkc1p, the yeast protein kinase C homolog, which is thought to regulate the Mpk1p MAP kinase pathway involved in cell wall remodeling and polarized cell growth, was found to act as a dosage suppressor of *glc7-10*. Therefore, Glc7p is required for normal cell wall integrity, bud development and cell cycle progression. In addition, *glc7-10* cells require higher levels of Pkc1p activity to sustain these functions (Andrews and Stark, 2000).

Sequencing of the coding region of *glc7-10* revealed one point mutation at position +928, causing an F135L substitution, a residue that is conserved in all members of the PPP family of protein phosphatases (Andrews and Stark, 2000). The minimum restrictive temperature of *glc7-10* is ~37ºC.

1.8.2.2 glc7-12

The temperature-sensitive Glc7p allele, *glc7-12*, causes a block to the completion of mitosis at the restrictive temperature (MacKelvie et al., 1995). At 37ºC,

around 35% of *glc7-12* cells arrest with a nucleus stretched across the neck of the bud and a short mitotic spindle. Additional copies of SDS22, a known regulator of the mitotic function of PP1 in *Schizosaccharomyces pombe* (Ohkura et al., 1989) lead to the suppression of the *glc7-12* mutant (MacKelvie et al., 1995). This suggests that the interaction of the two proteins is of functional significance and indicates that Sds22p is a regulatory subunit of Glc7p during mitosis. However, even at high copy number, SDS22 was unable to rescue growth of either glc7-10 or *glc7-13*. This allele-specific effect of Sds22p indicates that the function of Sds22p is related to the role of PP1 during mitosis.

Sequencing of the coding region of *glc7-12* revealed three separate point mutations at positions +93, +734 and +1261. However, since the mutation +734 is silent, *glc7-12* encodes a product with two predicted amino acid changes: N32Y and G227D. The N32Y substitution is not highly conserved but it lies immediately adjacent to the motif [-E(I/V)R-] conserved in PP1. The central G227D substitution alters a residue highly conserved in PP1 which marks the boundary of a predicted α-helical region (MacKelvie et al., 1995). The minimum restrictive temperature of *glc7-12* is ~33-34ºC.

1.8.2.3 glc7-13

The temperature-sensitive *glc7-13* mutant strain exhibits a random arrest with respect to the cell cycle suggesting that *glc7-13* is not involved in cell cycle progression (Mike Stark, personal communication). In addition, the cell size of *glc7-13* mutant cells is reduced at the restrictive temperature compared to WT (Mike Stark, personal communication). Sequence analysis of the *glc7-13* coding region revealed two point mutations at positions +143 and at +1049. Therefore, *glc7-13* has two amino acid substitutions: Q48R and L175P (Andrews and Stark, 2000). The minimum restrictive temperature of *glc7-13* is ~32ºC.

1.8.2.4 glc7-5

Glc7p promotes mRNA export by facilitating association of Mex67p, a conserved heterodimeric export receptor, with mRNA (Gilbert and Guthrie, 2004). Npl3p is required for mRNA export in *S. cerevisiae* probably by acting as an adaptor

for Mex67p. Gilbert and co-workers could show that nuclear dephosphorylation of Npl3p *in vivo* is promoted by Glc7p and that this is required for mRNA export (Gilbert and Guthrie, 2004). The authors proposed that Mex67p-dependent mRNA export is regulated by Glc7 dependent dephosphorylation of Npl3p (Gilbert and Guthrie, 2004). For their studies Gilbert and co-workers used the *glc7-5* ts allele. This implicates that glc7-5 mutant strains accumulate mRNAs in the nucleus and also that phosphorylated Npl3p accumulates in these mutants.

glc7-5 was also shown to arrest in the G1 phase of the cell cycle (M. Stark, personal communication). *glc7-5* has a single point mutation at position 1252, which changes a threonine to an alanine at position 224. The minimum restrictive temperature of *glc7-5* is ~32ºC.

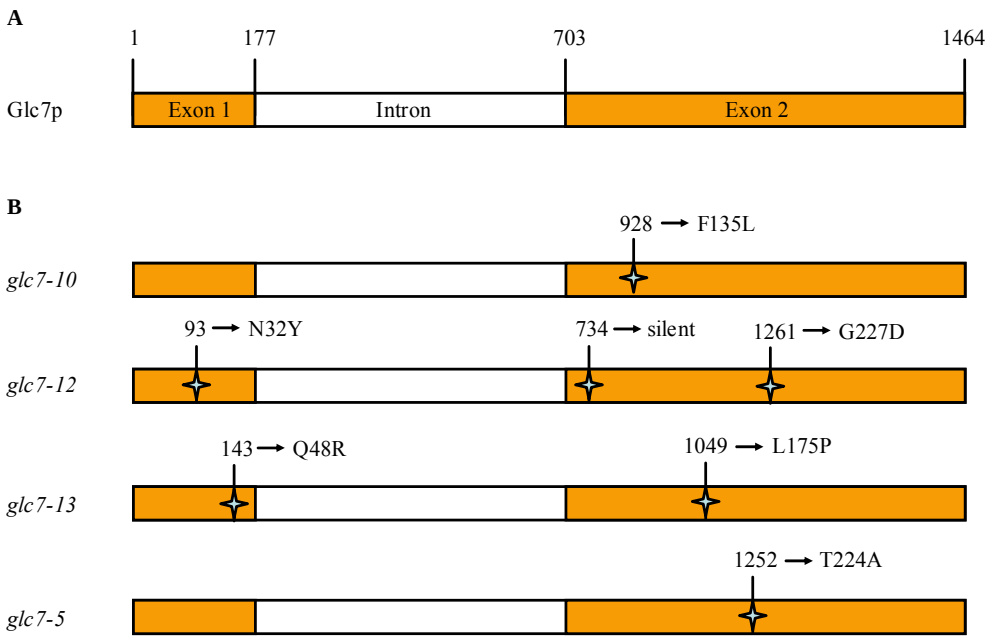


Fig. 1.6. Location of point mutations in *glc7* mutant alleles.
A) Schematic representation of the genomic organization of Glc7p.
B) Schematic representation of *glc7* mutants used in this study. Stars indicate the nucleotide positions of the point mutations that change the amino acid sequence.

2 Glc7p regulates polyadenylation but not cleavage of pre-mRNAs

2.1 Summary

Pre-mRNAs are cleaved and the upstream fragments are subsequently polyadenylated at their 3' end. Cleavage and polyadenylation factor (CPF) is a multiprotein complex conducting these two reactions in cooperation with cleavage factor IA (CF IA) and cleavage factor IB (CF IB). Here, we show that the protein phosphatase Glc7p, a component of CPF, is required for the polyadenylation step however not for the cleavage step of pre-mRNAs. We found that loss of Glc7p activity causes the accumulation of shortened poly(A) tails *in vivo*. GST-pulldown experiments revealed interactions with several subunits of CPF and CF IA. Moreover, the protein levels of several subunits of CPF are reduced in a *glc7* mutant strain, suggesting that CPF complex formation is compromised. Finally, we observe a poly(A) site selection defect in a *glc7* mutant strain.

These results indicate that Glc7p is essential for polyadenylation and poly(A) site selection of pre-mRNAs may be regulated via CPF complex assembly.

2.2 Introduction

pre-mRNAs of eukaryotes have to undergo extensive processing including addition of a 5' cap, splicing of introns and 3' end processing, to become mature, functional mRNAs. These processes all occur co-transcriptionally and are coupled to transcription via the CTD of RNAP II, which acts as a landing platform for processing factors (Proudfoot, 2000a).

3' end processing of pre-mRNAs consists of two steps: endonucleolytic cleavage at the poly(A) site followed by the addition of a poly(A) tail to the upstream cleavage product. Poly(A) tails are important for the stabilization, export and also for translation of mRNAs. Loss of the poly(A) tails leads to a rapid degradation of mRNAs. In *S. cerevisiae* the average poly(A) tail has a length of around 70 nucleotides. Cis-acting sequences in the 3' untranslated regions (3'UTRs) guide the trans-acting cleavage and polyadenylation factors to the correct cleavage and polyadenylation site. In *S. cerevisiae* the CPF, CF IA, CF IB and the poly(A) binding protein (Pab1p) are required for the cleavage and polyadenylation reactions *in vitro* (Chen and Moore, 1992; Ohnacker et al., 2000). CPF has been affinity purified from

yeast extracts (Ohnacker et al., 2000). Mass spectrometry of this purified CPF led to the identification of Ref2p, Pti1p, Swd2p, Ssu72p and Glc7p, formerly unknown subunits of CPF (Dichtl et al., 2002a). In addition, CPF contains the putative endonuclease Ysh1p, the poly(A) polymerase Pap1p and the subunits Yhh1p, Ydh1p, Yth1p, Pfs2p, Pta1p, Mpe1p (Ohnacker et al., 2000). Several other groups also found these subunits to be part of CPF (Gavin et al., 2002; Walsh et al., 2002). CF IA was shown to comprise the subunits Rna14p, Rna15p, Pcf11p and Clp1p (Kessler et al., 1996). CF IB consist of a single subunit Nab4p/Hrp1p (Kessler et al., 1996). Several subunits of CPF have been shown to be involved in recognition of the poly(A) site including Yhh1p, Ydh1p and Yth1p (Barabino et al., 2000; Dichtl et al., 2002b; Dichtl and Keller, 2001; Kyburz et al., 2003). These factors bind to sequences surrounding the poly(A) site, which specify the actual cleavage site *in vivo* (Dichtl and Keller, 2001).

Phosphorylation of proteins is a major control mechanism in the cell. It was proposed that 3' end formation is also regulated via phosphorylation events, since the polyadenylation activity of Pap1p is downregulated by phosphorylation *in vitro* (Mizrahi and Moore, 2000). However, this type of regulation seems to be restricted to the S/G2 transition stage of the cell cycle (Mizrahi and Moore, 2000). Therefore, Pap1p might not be the target of regulating polyadenylation at each round of transcription of a gene.

Glc7p is a serine/threonine phosphatase that has been implicated in many cellular processes in the nucleus and the cytoplasm. These include glycogen accumulation, carbohydrate metabolism, translation initiation, transcription, mitosis, meiosis, vesicular trafficking, cell polarity, histone modification and ion homeostasis (Stark, 1996). Glc7p is the only yeast homolog of the mammalian protein phosphatase 1 (PP1) and their protein sequences are 81% identical (Feng et al., 1991). To achieve specificity for different substrates Glc7p associates with targeting or regulatory subunits that direct it to the substrate and/or regulate its activity. Thus, it is possible that a single phosphatase carries out many distinct reactions. The association of Glc7p with CPF suggests that a targeting and/or regulatory subunit may recruit Glc7p to the complex. This could also imply that Glc7p is involved in the regulation of cleavage or polyadenylation by phosphorylation. While our studies on the function of Glc7p in 3' end processing were ongoing, He and co-workers reported that Glc7p is required for polyadenylation of pre-mRNAs (He and Moore, 2005). The authors demonstrated that

the phosphorylation state of Pta1p is regulated by Glc7p and that dephosphorylation of Pta1p is necessary to switch CPF to a poly(A) adding-competent complex.

Here, we show that glc7-13 mutants are defective for polyadenylation but not for cleavage of pre-mRNAs *in vitro* and *in vivo*. In addition, the *glc7-12* mutant strain displays a poly(A) site selection defect. Loss of Glc7p activity leads to reductions in protein levels of several subunits of CPF.

2.3 Results

2.3.1 GST-pulldown experiments reveal physical interactions between CPF subunits and Glc7p

The *S. cerevisiae* protein phosphatase Glc7p co-purified with CPF (Dichtl et al., 2002b; Walsh et al., 2002). We therefore decided to identify interaction partners of Glc7p within CPF. To this end we performed GST-pulldown experiments. Glc7p was expressed as a fusion protein with glutathione-S transferase (GST) and incubated with *in vitro* translated subunits of CPF and CF IA. As a control GST alone was incubated with the *in vitro* translated proteins to test for background binding of GST to these proteins. We observed strong interactions of Glc7p with Pfs2p, Ysh1, Pta1p, Pap1p, Pcf11p and Ssu72p. Intermediate interactions are observed for Yhh1p, Ydh1p, YDL094c, Clp1p and Glc7p itself. Weak GST-pulldown interactions were detected for Rna14p. No GST-pulldown interactions with Glc7p were observed for Pti1p, Ref2p, Mpe1p, Fip1p, Pab1p, Rna15 and Nab4p (Fig. 2.1 a; Table 2.1). All these numerous interactions of Glc7p with CPF indicate that Glc7p is an integral part of CPF. Moreover, interactions designated as "strong" could have identified Ysh1p, Pap1p, Pta1p, Pfs2p, Ssu72p and Pcf11p as the minimum core interaction partners of Glc7p within CPF and CF IA.

It is striking that these core interactions include the putative endonuclease Ysh1p and the poly(A) polymerase Pap1p, as this may indicate an important function for Glc7p in the regulation of 3' end formation.

Correlating with our results, it was reported recently that Pta1p is a phosphoprotein that is targeted by Glc7p to regulate polyadenylation of pre-mRNAs (He and Moore, 2005). These data also strengthen the idea that the core interactions we observed in our *in vitro* GST- pulldown assays reflect the situation *in vivo*.

We do not observe an interaction of Glc7p with Ref2p in our pulldown experiment. Contrary to our observation, an interaction between the two proteins was demonstrated by yeast two hybrid experiments and by affinity purification (Nedea et al., 2003; Uetz et al., 2000).

We also tested whether GST-Glc7p interacts with proteins involved in transcription or processing of pre-mRNAs. Those included Rpb2 (second largest subunit of RNAP II), Sua7p, a general transcription factor required for transcription initiation and start site

selection by RNAP II, which interacts with Ssu72, Upf1p, an ATP-dependent RNA helicase of the SFI superfamily, required for nonsense mediated mRNA decay and for efficient translation termination at nonsense codons and Nrd1p, an RNA-binding protein that interacts with the C-terminal domain of the RNAP II large subunit and is required for transcription termination and 3' end maturation of non- polyadenylated RNAs. GST-Glc7p showed strong interaction with Rpb2p, Sua7p and Upf1p and moderate binding with Nrd1p (Fig. 2.1 b, Table 2.1), suggesting further functions of Glc7p in the transcription process.

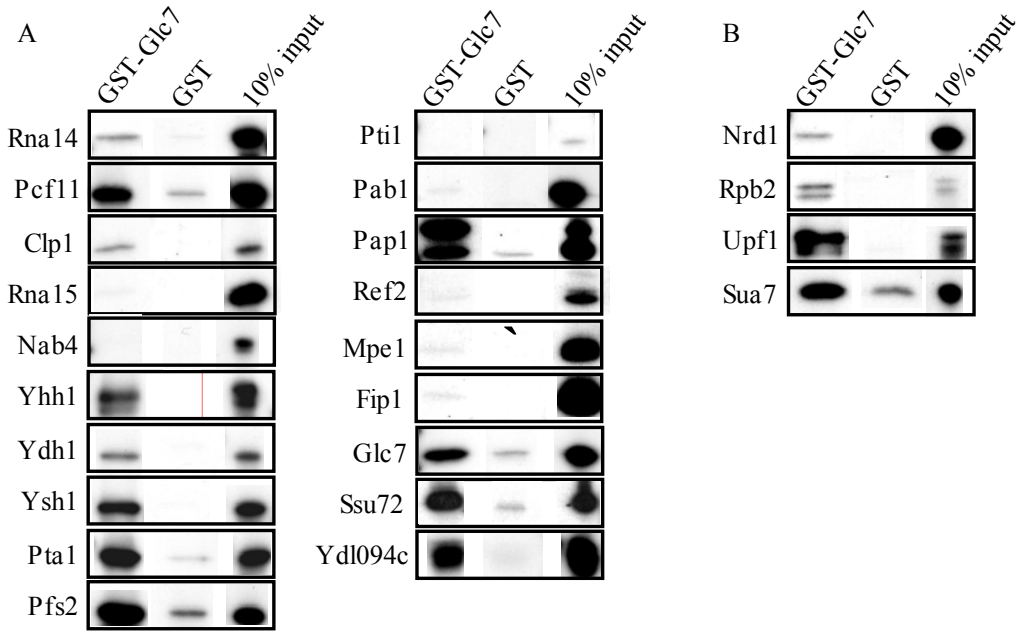


Figure 2.1. Glc7p interacts directly with subunits of CPF and CFIA
(A) GST (lane 2) and GST-Glc7p (lane 1) bound to Glutathione Sepharose 4B were incubated with in vitro translated ^{35}S-labeled subunits of CPF and CFIA as indicated on the left of the panel. Bound proteins were separated by SDS-PAGE and visualized by autoradiography. Input (lane 3) shows 10% of the in vitro translation reaction included into the binding reaction.

Table 2.1

strong interaction	medium interaction	weak interaction	no interaction
Pfs2	Yhh1	Rna14	Pti1
Pta1	Ydh1	Nrd1	Ref2
Ysh1	Glc7		Mpe1
Pap1	Ydl094c		Fip1
Pcf11	Clp1		Pab1
Rpb2			Rna15
Upf1			Nab4
Sua7			
Ssu72			

2.3.2 The glc7-13 mutant strain shows a defect in polyadenylation but not cleavage of pre-mRNAs in vitro

3' end processing of pre-mRNAs is a two step reaction: endonucleolytic cleavage followed by polyadenylation of the upstream cleavage product (Zhao et al., 1999a). Subunits of CPF are either involved in the cleavage or the polyadenylation reaction or in both. Since Glc7p co-purified with CPF and physically interacts with several subunits of CPF in GST-pulldown experiments, we wanted to gain more insight into the role of Glc7p in 3' end processing of pre-mRNAs. Therefore, we carried out *in vitro* cleavage and polyadenylation assays with protein extracts from WT and Glc7p mutant strains to asses whether Glc7p could have a function in one of these reactions. To this end, we prepared protein extracts of WT, *glc7-5*, *glc7-12*, *glc7-13* and *glc7-10* mutant strains. In order to find out whether cleavage is affected we tested these *glc7* mutants for their ability to cleave a CYC1 precursor RNA (CYC1 pre) *in vitro* at the permissive or non-permissive temperature. None of the *glc7* alleles showed a defect in cleaving the CYC1 pre-mRNA substrate at the restrictive temperature of 37º C (Fig. 2.2 a and b), indicating that Glc7p is not involved in the cleavage reaction of pre-mRNAs. However, in the polyadenylation assay, extracts prepared from *glc7-13* cells were not capable to polyadenylate a pre-cleaved CYC1 precursor RNA at 37º C (Fig. 2.2 a, Monika Garras, unpublished). In contrast, *glc7-12* and *glc7-5* mutant extracts showed no defect in polyadenylation, suggesting that the defect in polyadenylation is allele-specific.

We conclude that Glc7p is needed for the polyadenylation step of 3' end processing but not for the cleavage step.

This result is in agreement with recent data published by He and co-workers showing that a conditional null allele of Glc7p no longer polyadenylates an *in vitro* transcribed precursor RNA (He and Moore, 2005).

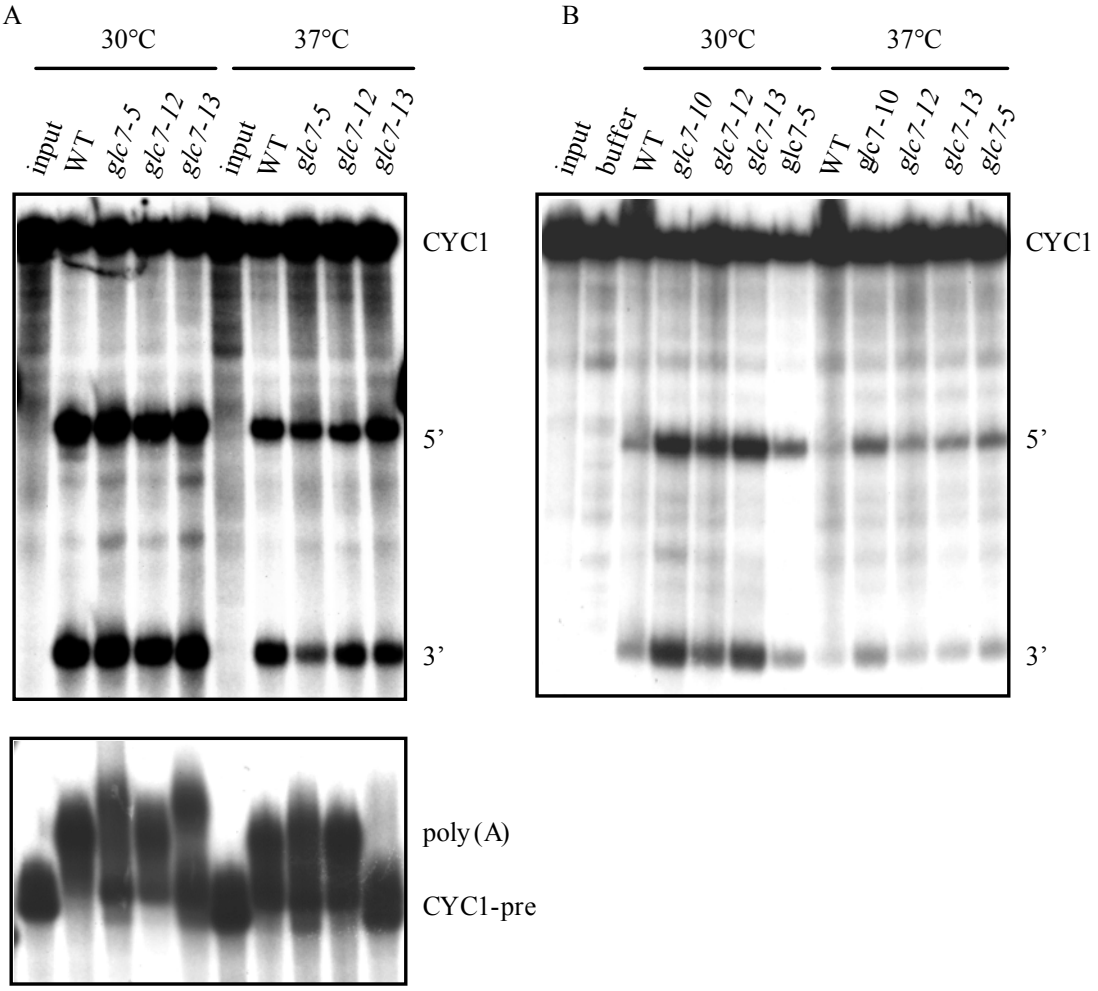


Figure 2.2. *glc7-13* is deficient in polyadenylation but not in cleavage of pre-mRNAs.
In vitro analysis of cleavage and polyadenylation activities of *glc7* mutant extracts (A and B). Input lanes represent mock-treated reactions. The migration positions of the CYC1 and CYC1-precleaved (CYC1-pre) RNA substrate, specific cleavage products (5' and 3') and the polyadenylation products (poly(A) are indicated on the right of each panel (A and B). Extracts prepared from yeast strains as indicated on top were monitored for their ability to cleave internally ^{32}P-labeled CYC1 RNA substrate at 30º C and 37º C. Specific polyadenylation was analysed at 30º C and 37º C with internally ^{32}P-labeled precleaved CYC1 RNA substrate that ends at the natural poly(A) site.

2.3.3 Poly(A) tails are shorter in glc7 mutants in vivo

To further substantiate the function of Glc7p in the polyadenylation reaction we were interested whether Glc7p was also important for poly(A) tail synthesis *in vivo*. Analysis of the global poly(A) tail length distribution of a cell's mRNA is helpful to identify a defect in polyadenylation *in vivo* and to get a general overview if gene expression might be altered in a given strain.

To investigate the global poly(A) tail length distribution *in vivo*, total RNA was extracted from WT or *glc7-10*, *glc7-12*, *glc7-13* and *glc7-5* mutant strains grown at RT or shifted to 37ºC for one or two hours. Tail lengths were determined by labeling the RNAs at their 3' ends, followed by RNase A treatment that leaves only the poly(A) tails intact. In this assay, we observed that poly(A) tails were significantly shorter in *glc7-12* and to a lesser extent also in *glc7-13* and *glc7-5* mutant strains compared to the WT, suggesting a function for Glc7p in polyadenylation of pre-mRNAs (Fig. 2.3). In contrast, poly(A) tails in the *glc7*-10 mutant strain were not affected. We conclude from these findings that Glc7p is necessary for poly(A) addition not only *in vitro* but also *in vivo*. In addition we propose that gene expression is partially compromised in *glc7-12* mutants.

After we have performed this experiment it was reported that a conditional null allele of Glc7p showed a marked decrease in the length of poly(A) tails *in vivo* (He and Moore, 2005), confirming our result.

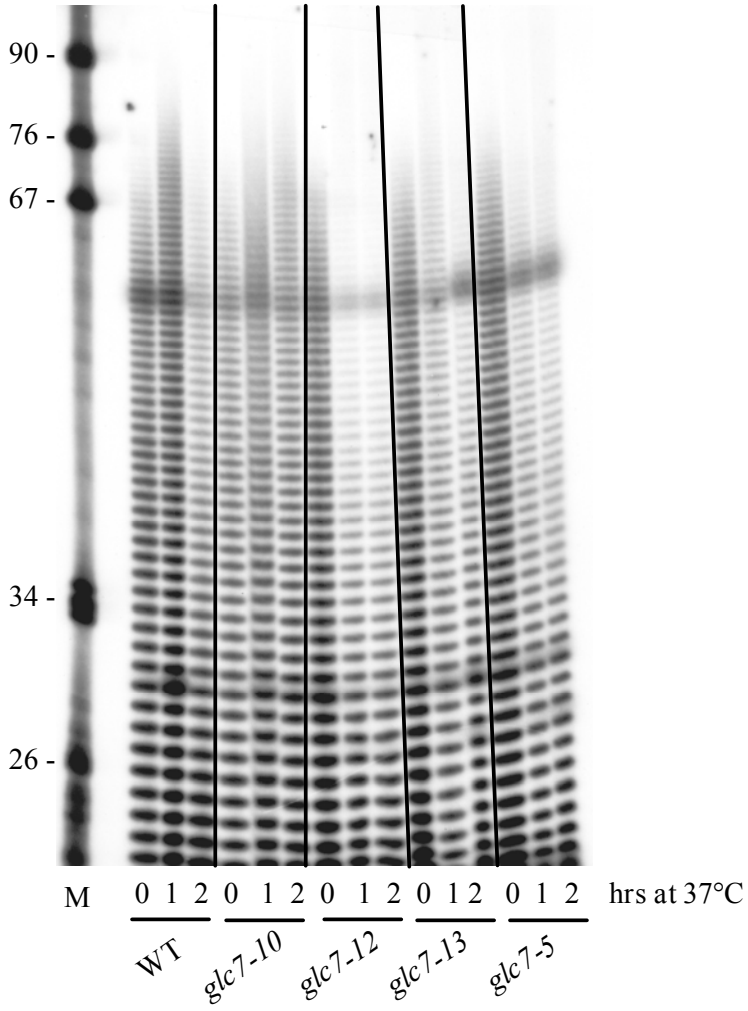


Figure 2.3. Poly(A) tail length distribution is affected in *glc7-12* mutant strains
Poly(A) tail length analysis of total RNA extracted from WT, *glc7-10*, *glc7-12*, *glc7-13* and *glc7-5* mutant strains after growth at 25º C and 37º C for various time points as indicated at the bottom of the panel. The position and size (in number of nucleotides) of the marker bands are indicated on the left.

2.3.4 glc7-13 mutant cells show a poly(A) site selection defect

To complete our study on the function of Glc7p in polyadenylation, we were interested if Glc7p might also be involved in selection of the poly(A) site. 3' untranslated regions (UTRs) of mRNAs can contain several poly(A) sites. In most cases, however, the most proximal site is the main poly(A) site which is predominantly used. A well known example of a gene that encodes multiple poly(A) sites is ACT1 (Mandart and Parker, 1995). The ACT1 3' UTR contains at least four potential polyadenylation sites. WT cells mainly use the most proximal poly(A) site and to a lesser extent downstream sites.

To test for poly(A) site selection, we carried out Northern analysis of WT and *glc7-10*, *glc7-12*, *glc7-13* and *glc7-5* mutant strains with a probe directed against the ACT1 mRNA. This analysis allows detection of the most proximal poly(A) site (site I) and the most distal poly(A) site (site IV) of ACT1. When cells were shifted to the restrictive temperature of 37˚ C for two hours, an increased signal for poly(A) site IV can be observed in *glc7-13* mutants, indicating that site IV is now as frequently used as site I in this mutant (Fig. 2.4). The other *glc7* mutants did not show an alteration in the selection of the poly(A) site suggesting that this phenotype is allele-specific.
We therefore propose that Glc7p is involved in poly(A) site selection of ACT1 pre-mRNA.

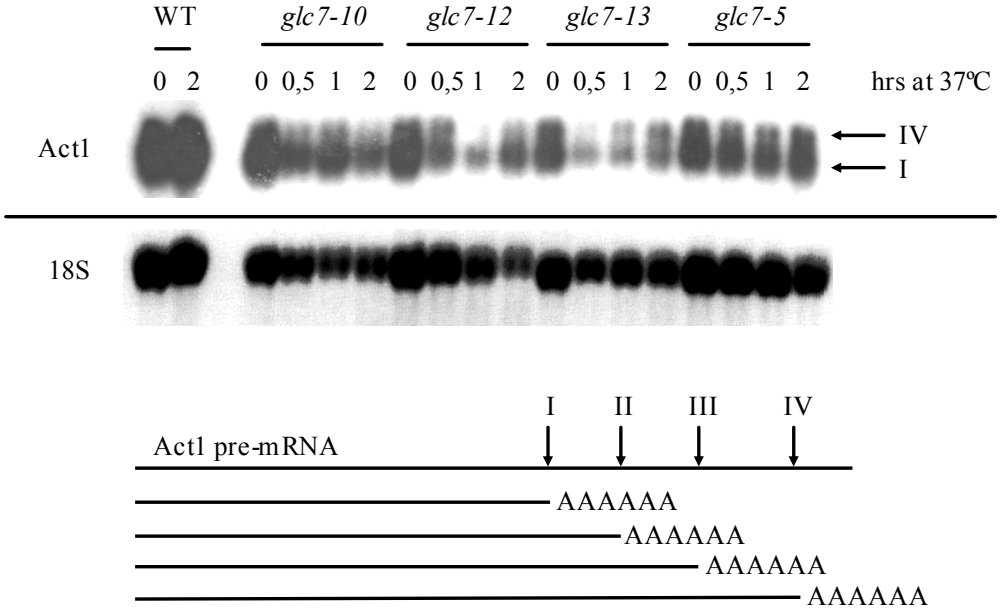


Figure 2.4. *glc7-13* mutant strains are deficient in recognition of the ACT1 poly(A) site *in vivo*
The drawing below the figure shows the relative positions of the different poly(A) sites. Northern analysis of total RNA extracted from WT and *glc7-10*, *glc7-12*, *glc7-13* and *glc7-5* mutant strains grown at 25º C or after shift to 37º C for 0.5, 1 and 2 hours as indicated at the top. Bands corresponding to site I or IV as indicated on the right. Blot was proped against Act1 and 18S, which served as a loading control.

2.3.5 The protein levels of several CPF subunits are altered in glc7 mutants

One possible explanation for the defects observed with some *glc7* mutants in the previous chapters is that CPF did not assemble correctly. The fact that Glc7p is part of the CPF complex rises the question whether mutations in Glc7p could destroy the

complex and consequently its function in cleavage and polyadenylation. Reduced activity of the phosphatase Glc7p might lead to a destabilization of the interactions between subunits of CPF since modifications of these subunits could be important for the association with the complex or for their function. Therefore the observed defects in polyadenylation in *glc7* mutants could be the result of a loss of CPF subunits from the complex that have important functions in polyadenylation.

To assess protein levels of distinct subunits of CPF in *glc7* mutant strains, we performed Western analysis on protein extracts derived from WT, *glc7-10*, *glc7-12*, *glc7-13* and *glc7-5* mutant cells before and after a one hour shift to the restrictive temperature (37˚ C) using antibodies against several subunits of CPF and CF IA (Fig. 2.5).

Results of the Western analysis showed a significant decrease of protein levels in the CPF subunits Fip1p, Pfs2p, Ysh1p, Pti1p, Pta1p and a modest decrease of Pap1p levels in the mutant strain *glc7-12* at the restrictive temperature. All of the CF IA subunits tested (Rna15p, Rna14p, Pcf11p) only showed a slight decrease in protein levels in *glc7-12* mutants. In contrast, none of the other *glc7* mutants resulted in a reduction in protein levels within these subunits. These data again suggests that there seems to be an allele-specific separation of phenotypes. Our results are only in part in accordance with those of He and co-workers (He and Moore, 2005). These authors also observed a decrease in protein levels of Pta1p and Fip1p, however not for Rna15p and Pap1p. This difference could be a result of the different strains used in these two studies. We were working with temperature-sensitive (ts) alleles whereas He and co-workers used a conditional null allele of Glc7p.

In conclusion, mutations in Glc7p seem to destabilize subunits of CPF and CF IA. This destabilization could contribute to the defects observed for polyadenylation in *glc7-12 in vivo*.

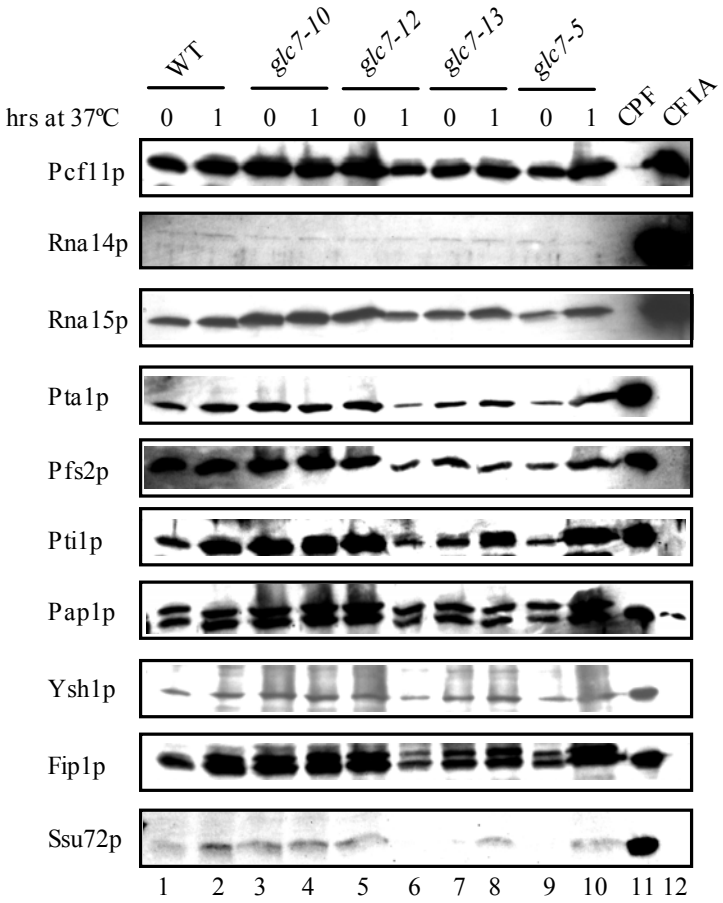


Figure 2.5. The amount of CPF subunits is reduced in *glc7-12* mutant strains
Western analysis of WT and *glc7* mutant extracts prepared from cells grown at 25º C, or after shift to 37º C for 1 hour as indicated on top of the panel. Equal amounts of total protein were loaded in each lane. Lane 11 shows purified CPF. Lane 12 shows purified CF IA. The filters were treated with antibodies directed against the proteins indicated at the left.

2.4 Discussion

Phosphorylation of proteins regulates many cellular processes. The essential protein phosphatase Glc7p, the yeast homolog of mammalian PP1, has been affinity purified together with CPF (Dichtl et al., 2002a; Walsh et al., 2002). This suggested to us that Glc7p could have a function in 3' end processing of pre-mRNAs. Here, we show that polyadenylation of a pre-cleaved RNA substrate is impaired in an *in vitro* cleavage and polyadenylation assays with whole cell extracts prepared from several *glc7* mutant strains. The involvement of Glc7p in the polyadenylation reaction was further supported by the fact that Glc7p also regulates poly(A) tail synthesis *in vivo* as poly(A) tails in several *glc7* mutant strains were shorter than poly(A) tails in WT cells. In addition, in *glc7-13* mutants the poly(A) site selected for 3' end processing shifted from the predominant WT poly(A) site towards a more distant cryptic poly(A) site. Western analysis of CPF subunits suggested that Glc7p phosphatase activity might regulate the assembly and therefore the stability of the CPF complex.

How could phosphorylation regulate CPF activity

One way to control the polyadenylation step of pre-mRNAs would be to directly regulate the activity of Pap1p by phosphorylation. Several reports indicate that Pap1 indeed can be phosphorylated and that this phosphorylation inhibits the enzyme resulting in the lack of poly(A) tails (Colgan et al., 1998; Mizrahi and Moore, 2000). However, phosphorylated forms of Pap1 have only be found in specific phases during the cell cycle: the late S/G(2) phase in yeast and the M phase in vertebrates (Colgan et al., 1998; Mizrahi and Moore, 2000). Therefore, this form of regulation does not seem to be involved in controlling the activity of Pap1 in each round of transcription. Moreover, in Xenopus and Hela cells it was demonstrated that Pap1 is heavily phosphorylated in a C-terminal region not conserved in the yeast enzyme (Colgan et al., 1998) suggesting that the mechanism of regulation by phosphorylation is quite diverse among species. Also, He and co-workers could show that in Glc7p depleted extracts Pap1p does not accumulate in a phosphorylated and therefore inactive form, indicating that Glc7p is not involved in the dephosphorylation of Pap1p (He and Moore, 2005).

In addition, these authors reported that Glc7p influences polyadenylation of pre-mRNAs by regulating the phosphorylation state of the CPF subunit Pta1p (He and Moore, 2005), which is phosphorylated by a unknown kinase. The phosphorylated state of Pta1p inhibits poly(A) tail synthesis. Glc7p dephosphorylates Pta1p at the end of each round of transcription leading to efficient polyadenylation of pre-mRNAs (He and Moore, 2005). However, it remains unclear how phosphorylated Pta1p inhibits Pap1p activity or how dephosphorylated Pta1p stimulates Pap1p activity. Phosphorylated Pta1p might destabilize CPF such that not all subunits required for polyadenylation are present during the elongation step of transcription. This could also be an efficient mechanism to prevent 3' end processing at cryptic poly(A) sites. Dephosphorylation of Pta1p during 3' end formation could also stimulate contacts to the missing subunits. This would lead to the formation of a stable CPF complex which can then efficiently polyadenylate the pre-mRNA. This effect of Pta1p might be mediated via the CPF subunit Fip1p, which helps to recruit Pap1p to CPF by interactions with another CPF subunit, Yth1p (Barabino et al., 2000; Zhelkovsky et al., 1998). Fip1p also regulates the processive activity of Pap1p (Preker et al., 1995). Pta1p could also be indirectly linked to Fip1p via interactions with subunits like Yth1p. Disturbing this interaction network could then exclude Fip1p from CPF and consequently Pap1p.

How is Glc7p recruited to CPF

Because Glc7p has many different functions at diverse locations in the cell, a specific targeting subunit has to recruit Glc7p to the CPF complex. It has been proposed that Ref2p is the CPF subunit that achieves the recruitment of Glc7p (Nedea et al., 2003). Nedea and co-workers demonstrated that when a specific region of Ref2p is deleted, Glc7p no longer co-purifies with the CPF complex (Nedea et al., 2003). This specific region of Ref2p contains a putative type I protein phosphatase binding site (amino acids 368-376), the canonical RVXF motif found in the majority of PP1/Glc7p interacting proteins (Walsh et al., 2002). Our pulldown experiments suggest that Glc7p interacts not only with Ref2p but with several subunits of CPF. These interactions on one hand seem to facilitate the recruitment and stabilize the interaction of Glc7p with CPF. On the other hand these interactions lead to an interaction of Glc7p with the substrate Pta1p.

3 Glc7p is involved in transcription termination of snoRNAs

3.1 Summary

The process that leads to the dissociation of RNA polymerase II (RNAPII) from the DNA template and the transcribed precursor RNA in the 3' UTR of a gene is called transcription termination. Transcription termination of pre-mRNAs and snoRNAs requires a number of cis- and trans-acting factors. Known components of snoRNA transcription termination are the Nrd1 complex and subunits of the CPF subcomplex APT. Here, we show that the protein phosphatase Glc7p is involved in transcription termination of snoRNAs but not of pre-mRNAs. Our data suggest that Glc7p functions in the Nrd1 complex-dependent pathway of snoRNA transcription termination. However, Glc7p does not regulate the phosphorylation status of any of the Nrd1 complex subunits. Also, phosphorylation of the APT subunit Pti1p and the CPF subunit Pta1p is unaltered in *glc7* mutant strains. Therefore it remains to be elucidated how Glc7p regulates snoRNA transcription termination.

3.2 Introduction

Two major classes of genes are transcribed by RNAP II: protein encoding genes and small nucleolar RNA (snoRNA) genes (Kiss, 2002). snoRNAs are non-coding RNAs that associate with a number of proteins to form small nucleolar RNPs (snoRNPs; (Kiss, 2002). The majority of snoRNAs functions as guide RNAs in the post-transcriptional synthesis of 2'-O-methylated nucleotides and pseudouridines in rRNAs, tRNAs and snRNAs and most likely other cellular RNAs (Kiss, 2002). Transcripts of both RNA classes have to be processed in order to become mature, functional RNAs. pre-mRNAs, which are the primary transcripts of protein encoding genes, are capped at the 5' end, spliced and 3' end processed. In contrast, the majority of snoRNA transcripts only undergo 3' end processing.

3' end processing of pre-mRNAs comprises two steps: endonucleolytic cleavage followed by the addition of a poly(A) tail to the upstream cleavage product. In yeast, this reaction requires the cleavage and polyadenylation factor (CPF), cleavage factor IA (CF IA) and cleavage factor IB (CF IB) (Chen and Moore, 1992; Ohnacker et al., 2000). CPF appears to form two distinct subcomplexes, the core-CPF subcomplex and the APT subcomplex (APT for associated with Pta1p). The core-CPF contains the

poly(A) polymerase Pap1p, the putative endonuclease Ysh1p, the subunits Ydh1p, Yth1p, Yhh1p, Fip1p, Mpe1p, Pfs2p and Pta1p. The APT subcomplex includes the six CPF subunits Ref2p, Pti1p, Swd2p, Ssu72p, Syc1p and Glc7p (Nedea et al., 2003). It has been shown that several of the APT subunits are involved in transcription termination of snoRNAs (Cheng et al., 2004; Dheur et al., 2003; Dichtl et al., 2004; Ganem et al., 2003; Steinmetz and Brow, 2003). In contrast, core-CPF subunits participate predominantly in processing of pre-mRNAs and not of snoRNAs.

Processing of snoRNAs can occur in several different ways. They all have in common that the snoRNA is cleaved within the 3' UTR and subsequently trimmed by the nuclear exosome to form the mature snoRNA (van Hoof et al., 2000). The nuclear exosome is a protein complex consisting of at least 10 different 3' to 5' exonucleases. Processing of snoRNA transcripts often depends on the endonuclease Rnt1p, which recognizes and cleaves specific hairpin structures in the 3' UTR thereby generating entry sites for the nuclear exosome (Chanfreau et al., 1997). However, Rnt1p cleavage sites are not present in all snoRNA precursors. Several snoRNAs contain sequences in their 3' region, which direct cleavage of the primary transcript by CF IA to generate the entry site for the nuclear exosome (Morlando et al., 2002).

Transcription termination is defined as the process of RNAP II dissociating from the DNA template and the nascent RNA. Transcription termination of protein coding genes is dependent on the 3' end processing signal, also termed poly(A) signal (Proudfoot, 2000a). Several factors of CF IA, including Pcf11p, Rna14p, Rna15p and the CPF subunits Ssu72p, Yhh1p, Swd2p and Ysh1p have been shown to be required for transcription termination of pre-mRNAs (Birse et al., 1998b; Dichtl et al., 2004; Dichtl et al., 2002b; Gross and Moore, 2001b; Minvielle-Sebastia et al., 1997; Steinmetz and Brow, 2003 ; M. Garras, personal communication). Similarly, transcription termination of snoRNAs requires cis- and trans-acting factors. The RNA-binding proteins Nrd1p and Nab3p and the putative RNA helicase Sen1p, all components of the Nrd1 complex have been implicated in transcription termination of snoRNAs (Steinmetz et al., 2001). A number of Nrd1 and Nab3 binding sites have been identified in snoRNA 3'UTRs and also in the Nrd1 ORF and 5' and 3' flanking regions (Carroll et al., 2004; Steinmetz et al., 2001; Steinmetz et al., 2006). They direct the termination of snoRNA transcription and also the autoregulation of the Nrd1 transcript itself (Arigo et al., 2006a; Vasiljeva and Buratowski, 2006). Nrd1p binds to the consensus sequence GUA[AG] whereas Nab3p binds to UCUU. The

arrangement of these motifs varies considerably at different snoRNA terminators, indicating a certain redundancy. In some snoRNAs, such as snR13, snR47 and snR65, the terminator region is arranged in a bipartite fashion (Steinmetz et al., 2006). Region I of these terminators contains the Nrd1p and Nab3p binding sites and shows some sequence similarity with the Nrd1 autoregulatory element. Interestingly, region II has sequence similarities with the cleavage and polyadenylation sites of pre-mRNAs and can direct polyadenylation, indicating that protein-coding and non-coding RNAP II transcribed genes might use similar mechanisms to direct transcription termination (Steinmetz et al., 2006). Previous studies have also described a requirement of the CTD kinase Ctk1p and the CTD of RNAP II itself for snoRNA transcription termination (Conrad et al., 2000; Steinmetz et al., 2001). Several studies suggest that the APT subunits Pti1p, Ref2p, Ssu72p, and Pta1p are participating in the Nrd1p-dependent pathway of snoRNA transcription termination (Dheur et al., 2003; Nedea et al., 2003). Recently, also the Paf1 complex which is involved in transcription elongation and histone modification, has been shown to take part in 3' end formation of snoRNAs (Sheldon et al., 2005). Furthermore, physical interactions of Nrd1p with the nuclear exosome have been reported, implicating the assembly of a 3' end processing/termination complex that couples transcription termination to processing of snoRNAs (Vasiljeva and Buratowski, 2006). However, the exact molecular mechanism of snoRNA transcription termination remains to be elucidated.

In this work we have analyzed the function of Glc7p in transcription termination of pre-mRNAs and snoRNAs. We show that Glc7p is essential for snoRNA but not for pre-mRNA transcription termination. Finally, we present data that Glc7p is not involved in the regulation of the phosphorylation status of Nrd1p, Sen1p and Nab3p, indicating that Glc7p does not influence snoRNA transcription termination via regulation of the Nrd1 complex.

3.3 Results

3.3.1 Expression profiling of glc7-12 reveals defects in snoRNA transcription termination

The CPF subunits Ssu72p, Pti1p, Ref2p and Swd2p (Dheur et al., 2003; Dichtl et al., 2004; Ganem et al., 2003; Steinmetz and Brow, 2003) have been shown to function in transcription termination of some snoRNAs. We therefore wondered whether Glc7p could also influence the transcription termination of these RNAs. To approach this question we performed an Affimetrix microarray analysis of poly(A)$^+$ RNA derived from *glc7-12,* a temperature sensitive Glc7p mutant, and the wild type strain at RT and 37º C to see if aberrant termination of snoRNAs affects the expression of downstream coding genes (for more detailed information concerning the microarray set up see Chapter 5 and Materials and Methods). A defect in transcription termination of snoRNAs can either lead to an increase in the RNA level of the downstream gene due to read-through or to a decrease resulting from interference with the downstream gene promoter. In several mutant strains (*ssu72-ts69, nrd1-5*), this approach has led to the identification of snoRNA transcription termination defects (Ganem et al., 2003; Steinmetz et al., 2001).

Results of our microarray analysis of *glc7-12* indicate nine mRNA genes with altered expression patterns that have a snoRNA located upstream (Table 3.1). The expression levels of Hem4, Trs31, Rim20, Rec104 and Yjl149w increased in *glc7-12* and are located downstream of the snoRNA genes snR5, snR13, snR31, snR71 and snR190-snR128, respectively. Steinmetz and co-workers (Steinmetz et al., 2001) found six instances in their microarray analysis of *nrd1-5* in which the mRNA levels downstream of snR3, snR13, snR47, snR71 and snR190-snR128 were increased. In addition, we found four mRNAs (Ytm1, Rpl7a, Ypl144w, Ckb2) with decreased expression levels. These mRNAs are located downstream of snR8, snR39b, snR17b and snR9, respectively. These results suggest an involvement of Glc7p in transcription termination of snoRNAs.

Interestingly, we also found downregulation of six mRNA genes encoding a snoRNA in their intron (Table 3.1). There are only seven protein coding genes in yeast that have a snoRNA in their intron. However, we did not further investigate these results.

Table 3.1.

ORF	upstream snoRNA	expression alteration
Hem4	snR5	upregulated
Trs31	snR13	upregulated
Rim20	snR31	upregulated
Rec104	snR71	upregulated
Yjl149w	snR190, snR128[a]	upregulated
Ytm1	snR8	downregulated
Ckb2	snR9[a]	downregulated
Ypl144w	snR17b[a]	downregulated
Rpl7a	snR39b	downregulated
Asc1	snR24[b]	downregulated
Tef4	snR38[b]	downregulated
Rpl17a	snR39[b]	downregulated
Rps22b	snR44[b]	downregulated
Imd4	snR54[b]	downregulated
Rpl17b	snR59[b]	downregulated

[a] ORF is followed by a snoRNA transcribed conversely

[b] snoRNA encoded in Intron of ORF

3.3.2 snoRNA read-through transcripts accumulate in glc7 mutants

Our microarray analysis of the *glc7-12* mutant strain suggested that a deficiency in snoRNA transcription termination is the cause for the up- or downregulation of genes downstream of a snoRNA. To test this we performed Northern analysis of total RNA extracted from WT, *glc7-10*, *glc7-12*, *glc7-13* and *glc7-5* strains grown at 25° C or after shift to the non-permissive temperature 37° C. In case of a failure to terminate transcription of a snoRNA, probes for a given mRNA that derived from the snoRNA downstream gene, should not only detect the respective mRNA but in addition a tandem transcript consisting of the snoRNA, the intergenic region and the downstream mRNA.

For this purpose probes directed against mRNAs (Trs31, Asn1, Rec104) that are located immediately downstream of snoRNAs (snR13, snR45, snR71) were used. Results for all probes showed the accumulation of extended RNAs (Fig. 3.1 a). The lengths of the extended RNA products correlated with RNA molecules that have been

demonstrated to consist of a snoRNA sequence followed by the 3' intergenic region and the sequence of the adjacent protein coding gene (Dichtl et al., 2002a; Ganem et al., 2003; Steinmetz et al., 2001). The amount of the read-through product varied significantly between the different *glc7* mutant alleles. *glc7-12* and *glc7-5* mutant strains displayed strong defects in terminating transcription at the snR13, snR45 and snR71 terminators. In contrast, *glc7-10* and *glc7-13* accumulated only very little read-through products within the same time interval at the restrictive temperature. Interestingly, in case of Trs31, only the read-through product is observed in the mutant strains. This indicates that ongoing transcription initiated from the snoRNA gene interfered with transcription initiation from the Trs31 promoter, leading to the downregulation of the Trs31 message itself.

If the read-through products we observed with the probe against the mRNA are indeed tandem transcripts, we can expect a probe against the upstream snoRNA to hybridize to the same read-through product. For technical reasons, attempts to obtain the same read-through products with probes directed against the snoRNA failed. We therefore used the method of primer extension on total RNA to demonstrate that transcription of this tandem transcript indeed started at the snoRNA transcription start site and fails to terminate at the snoRNA termination site. The method of primer extension allows to detect extended RNAs with oligonucleotide primers that are complementary to sequences in the intergenic region downstream of the termination site of snR13. In WT, the transcript does not contain the sequence complementary to this primer because transcription termination occurred upstream of the primer annealing site. Indeed, no reverse transcription product was observed in the WT strain (Fig. 3.1 b). In contrast, we observe a reverse transcription product in *glc7-5* mutant strains after shift to the non-permissive temperature (Fig. 3.1 b). This indicates that transcription did not terminate but continued at the snR13 termination site. Furthermore, the 5' end of this product coincides with the snR13 RNA 5' end. This result confirms that the read-through product, which we detected in the Northern analysis with the probe directed against the Trs31 mRNA was initiated from the snR13 transcription start site and eliminates the possibility that it initiated further 3' of the snoRNA transcription start site. Therefore, the read-through product we see in the Northern analysis is a tandem transcript consisting of the snoRNA, the intergenic region and the downstream mRNA. However, we did not observe a read-through product in *glc7-12* mutant strains, which was rather unexpected.

In conclusion, transcription termination at the snoRNA terminator did not occur but transcription continued into the downstream protein coding gene. This could suggests that *glc7* mutants have a defect in snoRNA transcription termination.

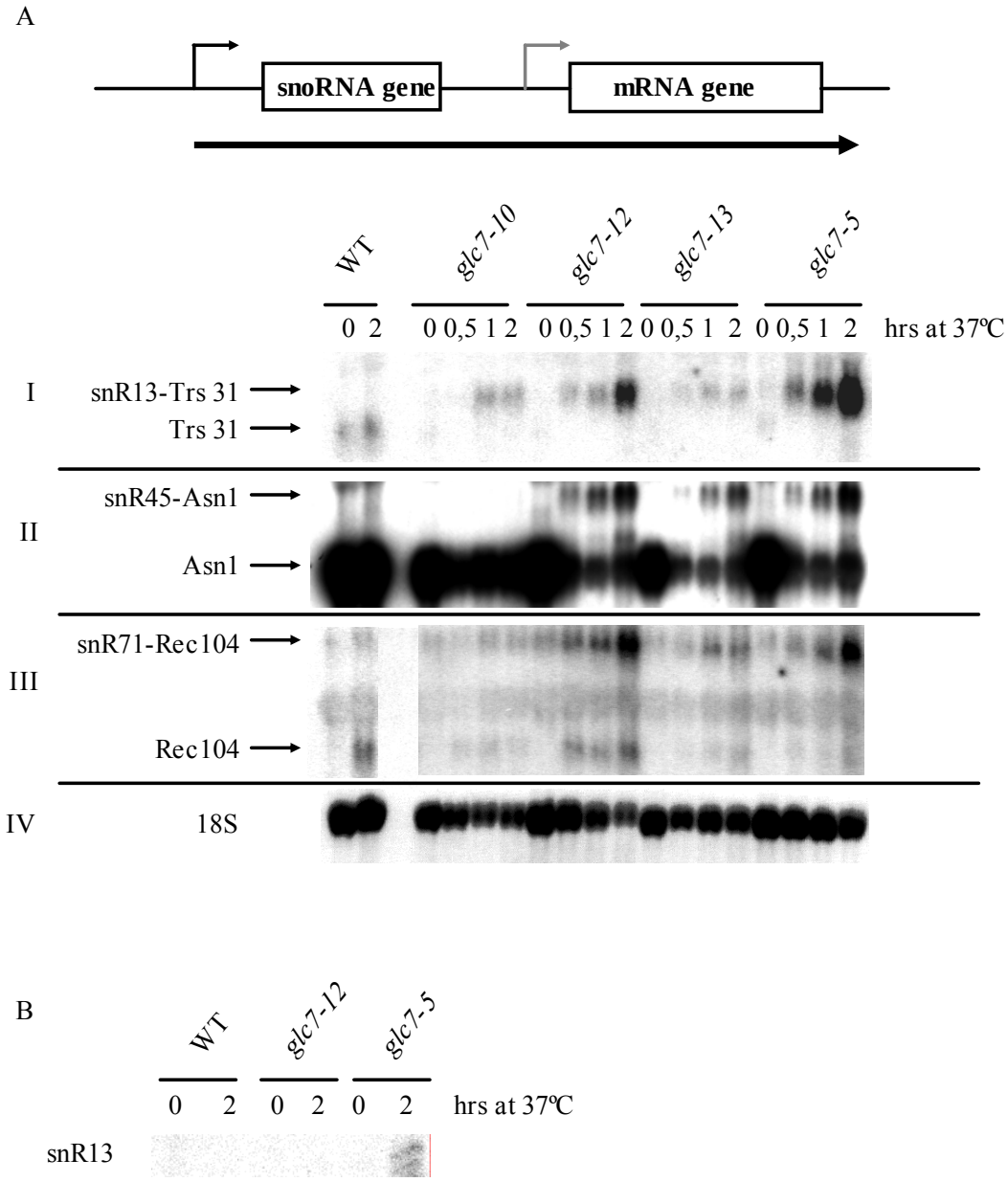


Figure 3.1. *glc7* mutant strains accumulate snoRNA-mRNA tandem transcripts.
The drawing on top illustrates the genomic arrangement of the analysed genes and the relative direction of transcription. A) Northern analysis (panel I – IV) of 20µg total RNA extracted from WT, *glc7-10*, *glc7-12*, *glc7-13* and *glc7-5* cells following growth in YPD medium at 25º C or after shift to 37º C for the indicated times. Blots were probed against Trs31 (panel I), Asn1 (panel II), Rec104 (panel III) and 18S (panel IV; loading control). Migration of RNAs is indicated on the left. B) Primer extension analysis of 6µg of total RNA extracted from WT, *glc7-12* and *glc7-5* cells following growth at 25º C or after shift to 37º C for 2 hours.

3.3.3 glc7-12 is defective in snoRNA transcription termination

A defect in transcription termination is not the only explanation for the accumulation of read-through products. Instead, it is possible that the accumulation of read-through products results from a stabilization of elongated transcripts that occurs exclusively in the mutants and not in the WT.

To discriminate between these possibilities and to establish a function of Glc7p in snoRNA transcription termination we performed transcriptional run-on analysis (TRO; see Materials and Methods). TRO monitors the distribution of actively transcribing RNAP II complexes over the genomic region of interest.

We tracked the density of RNAP II over the endogenous snR3 and the adjacent gene locus in the WT and *glc7-12* mutant strain *in vivo* as described by Steinmetz and co-workers (Steinmetz et al., 2001). The two single-stranded DNA probes A and B we used in the experiment, cover the genomic region over the snoRNA and the adjacent 5' end of gene YJR129c (see scheme in Fig. 3.2 a). These probes are complementary to transcripts initiated from the snR3 promoter. In WT, efficient transcription termination would result in run-on signals exclusively over probe A, since probe B is already complementary to regions past the snR3 termination site. Therefore, a signal over probe B indicates that correct transcription termination failed and transcription continued to the adjacent gene YJR129c. The gene YJR129c is transcribed in reverse orientation to snR3. As a result, mRNAs of YJR129c are not complementary to probe B and therefore cannot hybridize to probe B. Consequently, a run-on signal over probe B truly indicates that snR3 transcription termination has not occurred.

When testing the WT and the *glc7-12* mutant in this assay, we observed that *glc7-12* displayed an ~3 fold stronger run-on signal with probe B compared to the WT 45 min after shift to the non-permissive temperature, indicating that transcription termination was severely compromised in this mutant (Fig. 3.2 b and c). In addition, this result clearly shows that the read-through products detected in our northern analysis do not merely accumulate because trimming to the mature form failed. If RNAP II would read over the termination signal and consecutive degradation would lead to the formation of mature snoRNAs, we would expect a run-on signal over probe B for the WT as well. However, this is not the case.

In summary, these results again indicate that Glc7p is involved in transcription termination of snoRNAs.

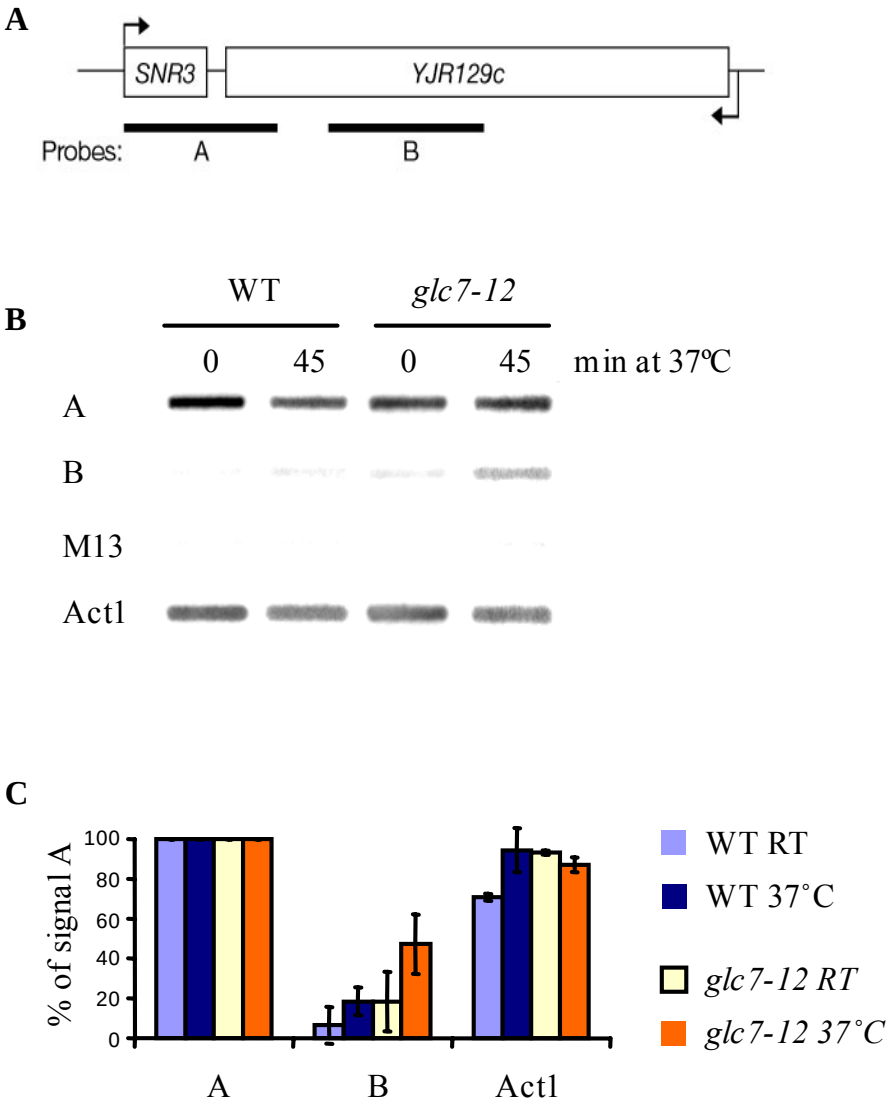


Figure 3.2. Glc7 is required for transcription termination of snoRNAs.
A) Arrangement of the SNR3 –YJR129C genomic locus. Probes (~360 nt each) are shown below their locations. Note the SNR3 and YJR129C are transcribed from opposite strands of the DNA, as indicated by the arrows. B) Representative slot blot from transcription run-on analysis from WT and *glc7-12* mutant cells before and after shift to 37º C. M13 slots are single-stranded phagemids with no insert and serve as a control for background hybridization. Transcription of Actin (Act1) serves as positive control. C) Quantification of the transcript by phosphorimager analysis. Background (M13) was substracted from each value and corrected counts were normalized to the A probe. Error bars represent standard deviation (n=3).

3.3.4 glc7 mutants accumulate Nrd1 and read-through transcripts at the Nrd1 transcription site

The yeast Nrd1 complex is required for transcription termination of RNAP II-transcribed snRNA and snoRNA genes (Steinmetz et al., 2001). In addition, two further phenotypes characterize mutations in the Nrd1 complex subunits Nrd1, Nab3

and Sen1. First of all, they lead to the accumulation of a tandem transcript (Nrd1-Mrpl17) composed of Nrd1, the intergenic region and the downstream mRNA Mrpl17, indicating that transcription termination at the Nrd1 terminator was inefficient. The reason for this is most likely the presence of Nrd1 and Nab3 binding sites in the 3' UTR of Nrd1, because these binding sites are also required for snoRNA transcription termination (Carroll et al., 2004). Second, the Nrd1 message itself accumulates in these mutants. It has been proposed that this phenotype is due to an auto-regulatory property of the Nrd1 complex that regulates the expression of the Nrd1 mRNA by premature transcription termination, a process that again involves Nrd1 and Nab3 binding sites, however in the 5'UTR and ORF of Nrd1 (Arigo et al., 2006a).

Considering these facts, we were interested if Glc7p is involved in the Nrd1-dependent pathway of transcription termination. If this was the case, *glc7* mutants should show the additional phenotypes of the Nrd1 complex mutants, such as accumulation of Nrd1 mRNA and the read-through product Nrd1-Mrpl17.

To this end, we performed northern analysis on total RNA extracted from WT, *glc7-10*, *glc7-12*, *glc7-13* and *glc7-5* mutant strains before and after shift to the non-permissive temperature. With probes directed against Nrd1 or Mrpl17, accumulation of the read-through product Nrd1-MRPL17 was detected for *glc7-12*, *glc7-13* and also to a lesser extent for *glc7-5* (Fig. 3.3). In contrast, strain *glc7-10* did not accumulate such a read-through product. In addition, all mutant strains showed an accumulation of the Nrd1 mRNA, which was most pronounced in *glc7-12* and *glc7-5*. An increase in the Nrd1 mRNA level at the restrictive temperature was also observed in *glc7-12* by microarray analysis. This implies that Glc7p also functions in regulation of the Nrd1 mRNA levels. Interestingly, *glc7-5* showed strong accumulation of Nrd1, similar to *glc7-12*, but did not accumulate the read-through product as efficiently as *glc7-12*. Therefore, these two phenotypes seem to be partly uncoupled in the *glc7-5* mutant.

We conclude that Glc7p is part of the Nrd1-dependent snoRNA transcription termination pathway.

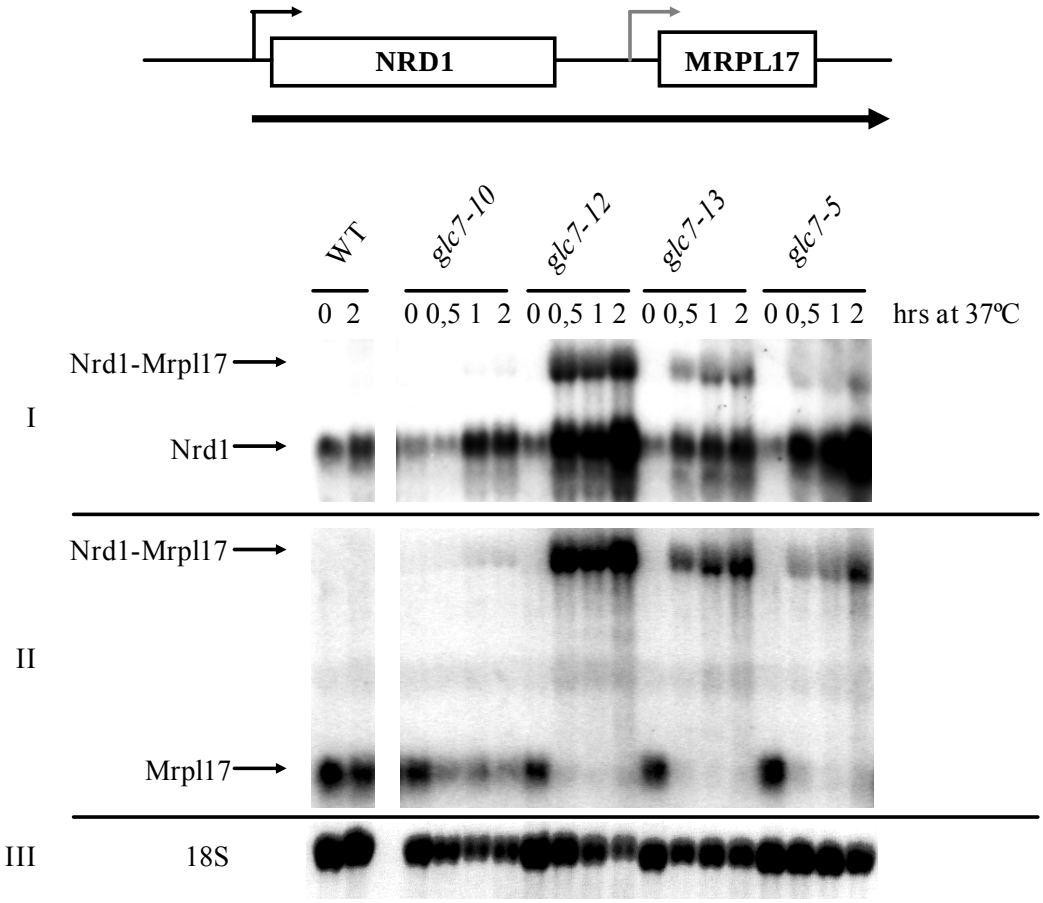


Figure 3.3. *glc7* mutant strains accumulate Nrd1-Mrpl17 tandem transcripts
The drawing on top illustrates the genomic arrangement of the NRD1-MRPL17 locus and the relative direction of transcription. Northern analysis (panel I – III) of 20µg total RNA extracted from WT, *glc7-10*, *glc7-12*, *glc7-13* and *glc7-5* cells following growth in YPD medium at 25º C or after shift to 37º C for the indicated times. Blots were probed against Nrd1 (panel I), Mrpl17 (panel II) and 18S (panel III; loading control). Migration of RNAs is indicated on the left.

3.3.5 glc7-12 does not show a transcription termination defect at the Cyc1 terminator

Read-through of the Nrd1 terminator as it appears in *glc7* and Nrd1 complex mutants, is a special case of transcription termination failure. Nrd1 is a protein-coding gene and its mRNA therefore is polyadenylated. The poly(A) signal in the 3' UTR is required for transcription termination of polyadenylated messages. However, Nrd1 has Nrd1 and Nab3 binding sites in addition to the poly(A) signal in the 3' UTR that are required for snoRNA transcription termination. We wondered whether the termination defect observed for the Nrd1 terminator in the *glc7* mutants was general, applying to all protein coding genes or whether it was due to the function of Glc7p in the Nrd1-

dependent pathway of snoRNA transcription termination. To answer this question we performed TRO analysis in *glc7-12*, monitoring the distribution of RNAP II with probes P1 to P6 that hybridize to the 3' end of the ORF and the following 3' UTR of the protein encoding gene Cyc1 *in vivo* (see scheme in Fig 3.4). The plasmid borne Cyc1 construct is a well documented substrate for TRO analysis of protein encoding genes (Dichtl et al., 2002a; Sadowski, 2003). In the WT, where transcription termination occurs shortly after the poly(A) site, run-on signals should be visible only over probe P1 to P3 but not over the probes located more downstream. A mutant strain defective for transcription termination would show additional run-on signals at probes P4 to P6.

We observed that compared to the WT, *glc7-12* did not show higher levels of actively transcribing RNAP II after the poly(A) site at the non-permissive temperature (Fig. 3.4). Therefore, transcription termination of Cyc1 is not affected in *glc7-12*, suggesting that read-through of the Nrd1 terminator in *glc7* mutant strains is associated with the functions of the Nrd1 complex and not with transcription termination defect of protein coding genes in general. However, the observation that *glc7-12* shows no transcription termination defect on the Cyc1 terminator could be allele- specific.

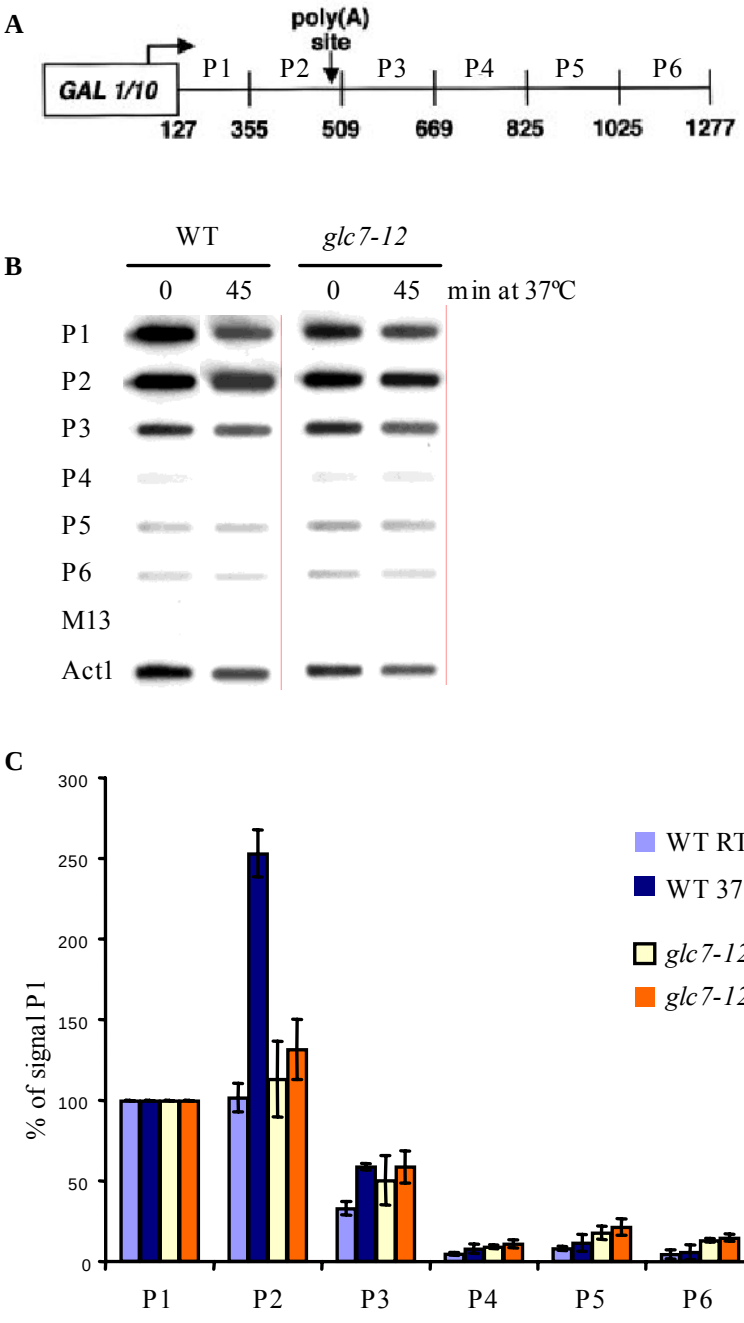


Figure 3.4. Glc7 is not required for transcription termination of CYC1.
A) Schematic diagram of plasmid pUGCYC1 showing the arrangement of M13 probes (P1 – P6) relative to the CYC1 poly(A) site (position 506). B) Representative slot blot from transcription run-on analysis from WT and *glc7-12* mutant cells before and after shift to 37º C for 45 minutes as indicated. M13 slots are single-stranded phagemids with no insert and serve as a control for background hybridization. Transcription of Actin (Act1) serves as positive control. C) Quantification of the transcript by phosphorimager analysis. Background (M13) was substracted from each value and corrected counts were normalized to the P1 probe. Error bars represent standard deviation (n=3).

3.3.6 glc7 mutants do not show a general transcription termination defect for protein encoding genes

To exclude the possibility that the terminating activity of Glc7p is restricted to the Cyc1 terminator and to a specific glc7 allele we carried out a transcriptional interference assay established by Morillon and co-workers (Morillon et al., 2003b). This assay allowed it to test a different mRNA terminator region, the Gal 10 terminator. Deficiencies in transcription termination could be recorded as follows (see Materials and Methods): transcription from the upstream pTEF:Kan^R fusion gene, if not terminated, will interfere with transcription initiation from the downstream Gal7 promoter (see scheme in Fig. 3.5 a). Because the Gal7 promoter drives expression of the GFP gene, interference with transcription from the Gal7 promoter will lead to a reduction in GFP signal, thus indicating a defect in transcription termination of a mutant strain.

As a positive control, we used the *yhh1-6* mutant strain. This strain was shown to have a transcription termination defect in a TRO experiment at the Cyc1 terminator (Dichtl et al., 2002b). *yhh1-6* exhibits a strong transcription termination defect in the transcriptional interference assay, confirming that our system is indeed functioning (Fig. 3.5 b). In contrast, *glc7-10*, *glc7-12* and *glc7-13* did not show a dramatic decrease of the GFP signal, indicating that transcription initiation from the Gal7 promoter is not disturbed by non-terminating polymerases coming from the TEF promoter. We observed a very low GFP signal in *glc7-5*, which probably does not correlate with a defect in transcription termination. We conclude this because the GFP signal is also very low in *glc7-5* mutants that bear a different construct from which transcription is only possible from the Gal7 promoter simply because the TEF promoter is deleted (delta Tef induced; see scheme in Fig. 3.5 a). With the help of this construct the transcriptional efficiency of GFP of a strain can be monitored. Because the GFP levels in *glc7-5* with the delta TEF construct were very low, we conclude that *glc7-5* has a deficiency in initiating or elongating transcription by RNAP II but not a transcription termination defect.

These results suggest that there is no transcription termination defect in *glc7* mutants for protein coding genes. In addition, we see that this effect is not allele-specific.

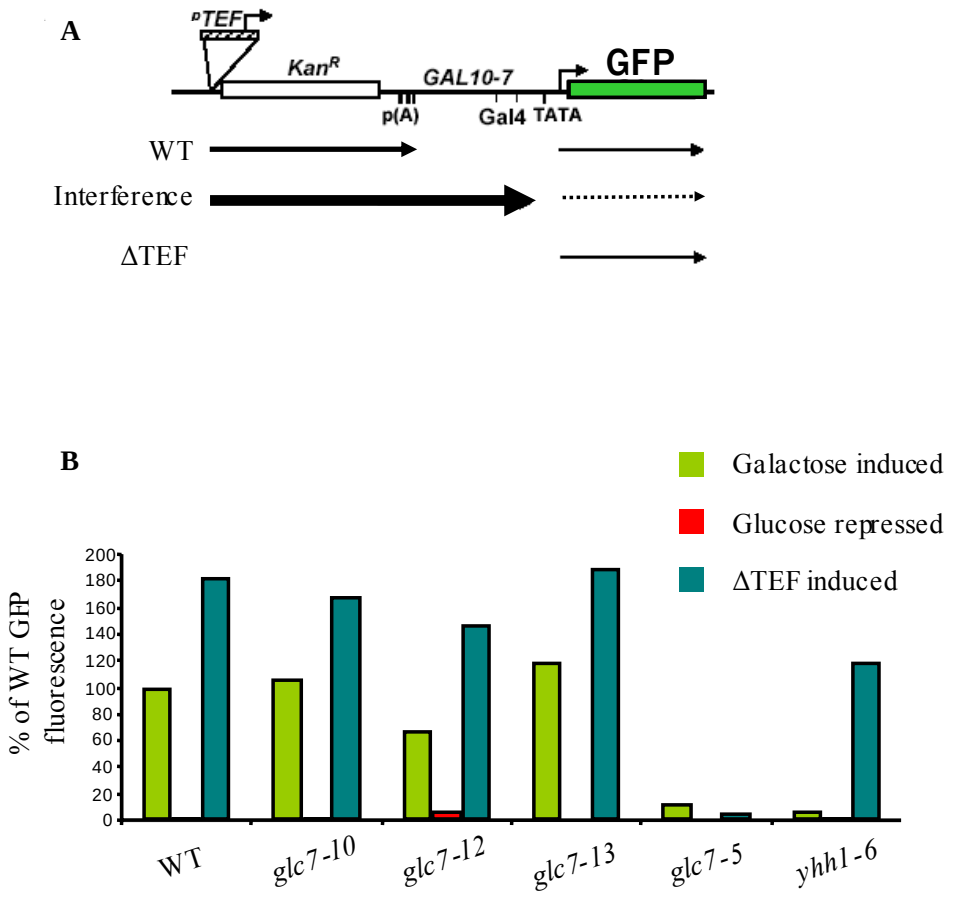


Figure 3.5. Glc7 is not required for transcription termination of GAL10.
A) Schematic of the transcription interference assay constructs. In yeast the GAL10-7 genes are expressed in tandem following GAL induction. Expression from the GAL7 promoter is very sensitive to transcriptional interference that occurs when transcription of the upstream GAL10 gene fails to terminate efficiently. A pTEF:Kan^R fusion gene is positioned upstream of the GAL10-7 intergenic region containing the GAL10 terminator and the galactose inducible GAL7 promoter driving expression of the GFP gene. Interference can be assessed by a reduction in GFP expression when the upstream pTEF:Kan^R fusion gene is expressed (pKGGURA). In order to control for transcriptional efficiency affecting the levels of GFP, a second version of the construct that lacks the TEF promoter, abolishing expression from Kan^R was transformed into strains (pKGGΔTEF). B) Plasmids pKGGURA and pKGGΔTEF were transformed into the strains indicated and assayed for GFP expression in synthetic medium containing glucose (represses expression from GAL7) or galactose (induced expression from GAL7) after shift to 37º C for 1 hour relative to WT. pTEF is expressed in both growth conditions. Expression of GFP (ΔTEF induced) serves as a control for transcription efficiency of the system.

3.3.7 Glc7p genetically interacts with subunits of CPF

It has been proposed that CPF can be subdivided into core-CPF and the APT sub complex, which stands for "associated with Pta1" (Nedea et al., 2003). The APT sub complex is composed of the subunits Pti1p, Ssu72p, Ref2p, Swd2p, Syc1p and Glc7p.

For all of these subunits, except for Syc1p, it was shown (including this work) that mutations in these protein factors lead to a snoRNA transcription termination defect (Dheur et al., 2003; Dichtl et al., 2004). The only exception is Ssu72p, which in addition shows a transcription termination defect for protein coding genes (Ganem et al., 2003; Steinmetz and Brow, 2003). Therefore, the APT subcomplex of CPF seems to be predominantly dedicated to the processing of snoRNAs.

To investigate a functional relationship between components of APT and Glc7p, we overexpressed several subunits in a 2µ vector in WT and *glc7-10*, *glc7-12*, *glc7-13* and *glc7-5* ts mutant strains. Overexpression of Pta1p, Pti1p and Syc1p dramatically inhibited growth of *glc7-12* at 33° C whereas growth of WT or *glc7-10* was not affected at 37° C (Fig. 3.6). Moreover, no effect was observed for *glc7-13* and *glc7-5* at any temperature. Overexpression of Ssu72p had no influence on growth of any of the *glc7* mutants. The only effect was seen on *glc7-12* with overexpressed Pta1p, Pti1p and Syc1p but not with Ssu72p. The other *glc7* mutant strains were not affected, indicating that this effect is allele-specific. Physical interactions between Glc7p, Pta1p and Pti1p have been described previously (Gavin et al., 2006; He and Moore, 2005; Nedea et al., 2003; Walsh et al., 2002). Because Pta1p and Pti1p are both phosphoproteins they are potential substrates of Glc7p in the context of snoRNA transcription termination.

Taken together, these experiments suggest that accumulation of Pta1p, Pti1 and Syc1p severely restricts growth in *glc7-12* cells at a semi-permissive temperature, illustrating that Glc7p regulates some aspects of the function of these proteins.

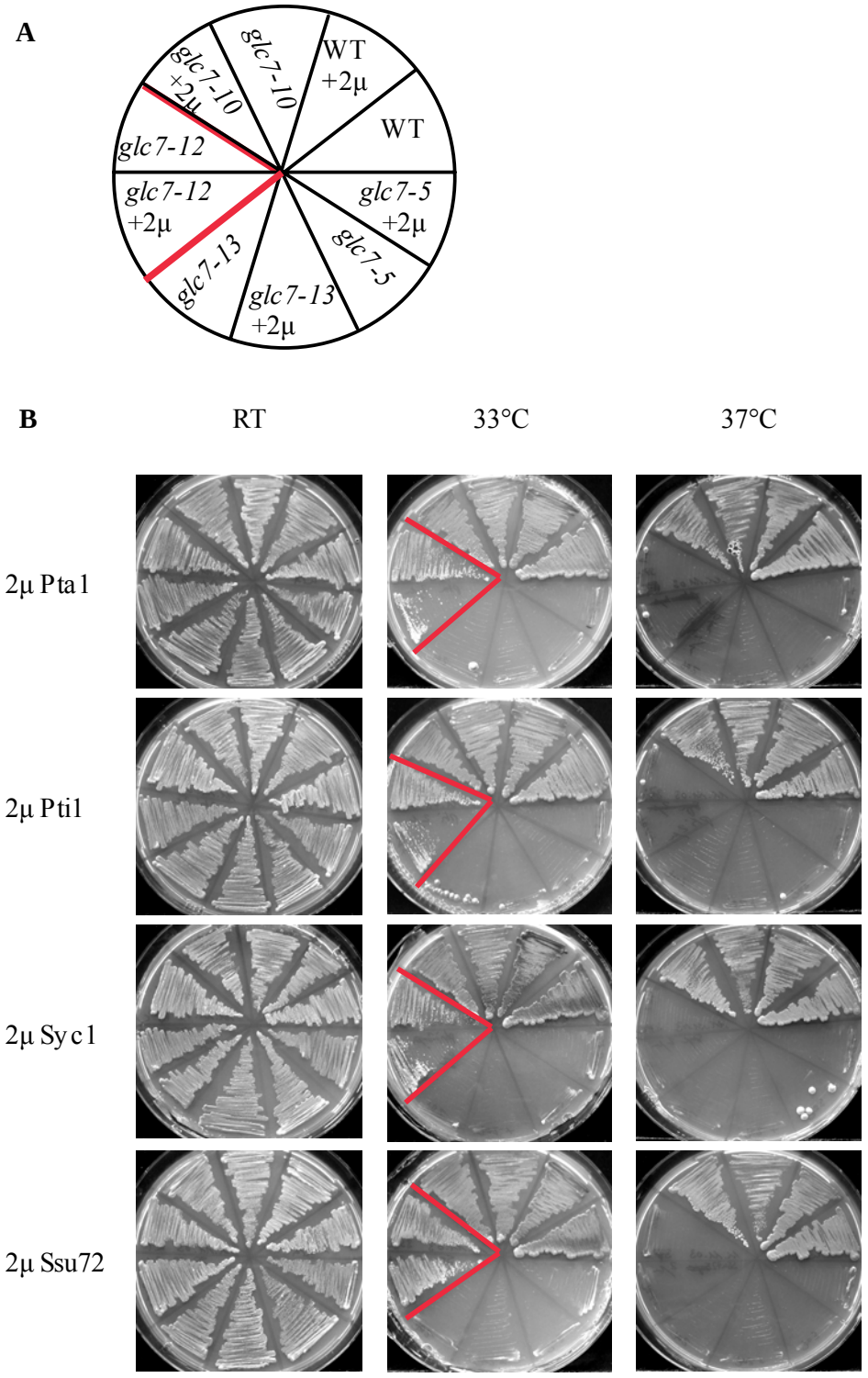


Figure 3.6. Overexpression of several APT subunits affects growth of *glc7-12* mutant cells.
A) Schematic representation of growth plate. B) WT and *glc7* mutant strains were transformed with 2µ overexpression vectors containing ORF of genes indicated on the left of each panel. Growth of strains was tested on synthetic medium at RT, 33º C and 37º C as indicated.

3.3.8 Pta1p is not the target of Glc7p in the process of snoRNA transcription termination

Recently, it was shown by He and co-workers (He and Moore, 2005), that phosphorylation of Pta1p is indeed regulated by Glc7p. These authors depleted Glc7p *in vivo* and observed that, in contrast to the WT, phosphorylated Pta1p accumulated in a Western analysis with an anti-phosphoserine antibody. Furthermore, it has been shown that phosphorylated Pta1p inhibits polyadenylation of pre-mRNAs, indicating that Glc7p regulates the polyadenylation activity of CPF via Pta1p. To test if Pta1p is also the target of Glc7p in the regulation of snoRNA transcription termination, we performed *in vitro* cleavage and polyadenylation assays with extracts purified from *glc7* mutants that showed snoRNA transcription termination defects. As depicted in Fig. 2.2 we found that the only *glc7* mutant allele to have a polyadenylation defect was *glc7-13*. Since *glc7-12* does not display an mRNA polyadenylation defect but is deficient in snoRNA transcription termination, we assume that regulation of snoRNA transcription termination does not involve regulation of Pta1p's phosphorylation status by Glc7p.

To confirm this, we tested if the phosphorylation level of Pta1p is affected in the *glc7-12* and *glc7-13* mutants compared to the WT. Therefore we affinity-purified TAP-tagged Pta1p from these mutants and the WT after shift to the restrictive temperature. We did not observe an alteration in the phosphorylation status of Pta1p in *glc7-12* mutants compared to the WT (data not shown). Unexpectedly, the phosphorylation state of Pta1p was not altered in *glc7-13* mutants as well.

We conclude that regulation of snoRNA transcription termination by Glc7p is not mediated by Pta1p.

3.3.9 Glc7p physically interacts with components of CF IA

3' end processing of pre-mRNAs is coupled to transcription termination (Birse et al., 1998b; Buratowski, 2005). The coupling is mediated by the poly(A) signal and the CTD (Fong and Bentley, 2001; Hirose and Manley, 1998; Howe, 2002; Zhang et al., 2005). Cleavage of the pre-mRNA is required but not sufficient for transcription termination to occur (Birse et al., 1998a; Kim et al., 2004b; Luo et al., 2006; Sadowski, 2003). Interestingly, CF IA, a complex required for the cleavage step in 3'

end processing of pre-mRNAs, has been shown to be involved in 3' end processing of snoRNAs (Morlando et al., 2002). Also, the CTD is necessary for 3' end processing/termination of snoRNAs. Consequently, there seems to be a certain similarity concerning the mechanisms of transcription termination for mRNAs and snoRNAs. We were interested, if Glc7p physically interacts with subunits of CF IA to answer the question whether Glc7p might influence snoRNA transcription termination via one of these factors. To this end, we carried out GST- pulldowns with GST-tagged Glc7p and *in vitro* translated subunits of CF IA. As shown in Fig. 3.7, Pcf11p has a strong affinity for Glc7p. Pcf11p is a protein that can bind to the CTD and was previously shown to dismantle the RNAP II elongation complex to support transcription termination of pre-mRNAs (Zhang et al., 2005). Rna14p and Clp1p made weak interactions and Rna15p did not interact with Glc7p.

In conclusion, Glc7p physically interacts with subunits of CF IA.

Since CF IA is not sufficient to carry out 3' end processing of snoRNAs on its own, but might interact with the Nrd1 complex, it is tempting to speculate that Glc7p might be involved in regulating this interaction.

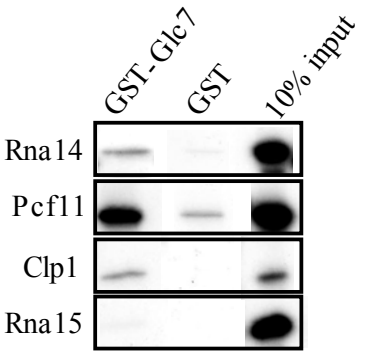


Figure 3.7. Glc7p interacts directly with subunits of CFIA (see also Fig. 2.1)

GST (lane 2) and GST-Glc7p (lane 1) bound to Glutathione Sepharose 4B were incubated with in vitro translated ^{35}S-labeled subunits of CPF and CFIA as indicated on the left of the panel. Bound proteins were separated by SDS-PAGE and visualized by autoradiography. Input (lane 3) shows 10% of the in vitro translation reaction included into the binding reaction.

3.3.10 Phosphorylation levels of the CTD are not altered in glc7 mutants

The CTD of RNAP II is required for transcription termination of snoRNAs. In yeast, the CTD consists of 26 repeats of the heptapeptide YSPTSPT, which can be

phosphorylated at serines 2 and 5. The dynamic pattern of phosphorylation of the CTD is important for binding of RNA processing factors including factors necessary for transcription termination. Considering the CTD binding property of Pcf11p and its strong interaction with Glc7p, we asked whether Glc7p might regulate the phosphorylation status of the CTD. Misregulation of the phosphorylation status could explain the transcription termination defects on snoRNAs observed in *glc7* mutant strains.

Total protein extracts prepared from the WT, *glc7-5, glc7-10*, *glc7-12*, and *glc7-13* mutants before and after shift to the non-permissive temperature were analyzed by Western analysis with antibodies directed against the unphosphorylated CTD (8WG16), against the CTD phosphorylated at Ser2 (H5) and against phosphorylation at Ser5 (H14). As shown in Fig. 3.8 we could not detect a change in the level of phosphorylation in these mutants compared to the WT. Therefore it is unlikely that the phosphorylation status of the CTD of RNAP II is regulated by Glc7p.

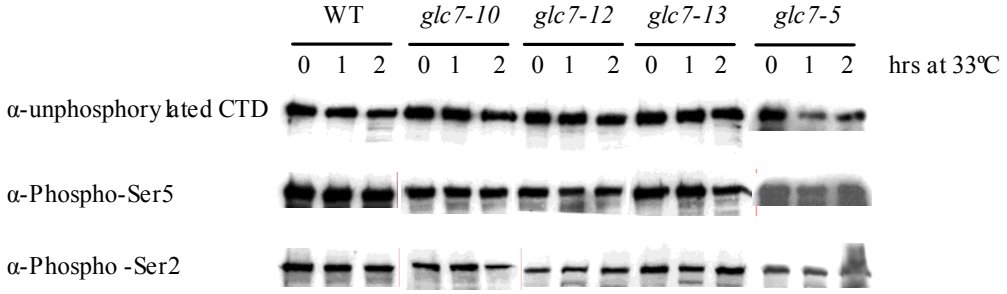


Figure 3.8. The phosphorylation level of the CTD is not altered in glc7 mutant strains.
Western analysis of WT and *glc7* mutant extracts prepared from cells grown at 25º C, or after shift to 33º C for 1 and 2 hours as indicated on top of the panel. Equal amounts of total protein were loaded in each lane. The filters were treated with antibodies directed against the unphosphorylated CTD (8WG16), the CTD phosphorylated at Ser5 (H14) and the CTD phosphorylated at Ser2 (H5) as indicated at the left.

3.3.11 Glc7p does not dephosphorylate any subunit of the Nrd1 complex in vivo

The Nrd1 complex subunits Nrd1p, Nab3p and Sen1p are all phospho-proteins. Because Glc7p is involved in the Nrd1-dependent pathway of snoRNA transcription termination it could be possible that Glc7p targets one of these proteins. We therefore decided to test whether phosphorylation of any of these proteins was regulated by the

phosphatase Glc7p. To this end, we performed Western analysis on protein extracts derived from WT, *glc7-12* and *glc7-13* mutant strains before and after shift to the non-permissive temperature. To detect phosphorylated forms of these proteins we used antibodies directed against phosphoserine and phosphothreonine residues. None of the Nrd1 complex subunits showed increased phosphorylation in the glc7 mutant strains at the restrictive temperature (B. Paguet, data not shown). This indicates that Glc7p is not involved in the regulation of phosphorylation of these proteins.

3.4 Discussion

Transcription termination is an important mechanism releasing the transcription machinery from the DNA template. Thereby RNAP II molecules are recycled for new rounds of transcription. In addition, transcription termination prevents transcription of downstream genes, whereas a failure to terminate transcription interferes with the proper expression of downstream genes. We report here that Glc7p, the yeast homolog of the mammalian protein phosphatase 1 (PP1) is required for transcription termination of snoRNAs, however not for transcription termination of pre-mRNAs. Transcriptional run on analysis revealed that snoRNA termination is deficient in a *glc7* mutant strain. This indicates that the cause for the accumulation of 3' extended snoRNAs is not a processing defect but a transcription termination defect. Furthermore, Glc7p showed all mutant phenotypes characteristic of mutations in factors of the Nrd1-dependent transcription termination pathway of snoRNAs, indicating that Glc7p has a function in this pathway. However, none of the Nrd1 complex subunits was found to be an *in vivo* substrate for Glc7p. Assaying the phosphorylation status of these proteins in several *glc7* mutant strains by Western analysis showed no Glc7 dependent change of phosphorylation.

Is it CPF-associated Glc7p that is involved in snoRNA transcription termination?

Glc7p has numerous functions in the nucleus as well as in the cytoplasm. Different targeting subunits that associate with Glc7p assure specificity for the reactions Glc7p is involved in by localizing or regulating the activity of Glc7p. It is difficult to prove whether Glc7p is involved in snoRNA transcription termination as part of CPF or as an independent unit. The problem is that the specific targeting subunit of Glc7 for this process is not known and in addition it is unclear whether the activity of Glc7p in this specific reaction depends on interactions with CPF. However, we assume that Glc7p as a subunit of CPF influences snoRNA transcription termination based on the following findings: Glc7p was found to be part of the CPF subcomplex APT. Five of the six APT subunits (including data of this work) have been shown to be required for transcription termination of snoRNAs (Dheur et al., 2003; Dichtl et al., 2002a; Nedea et al., 2003; Steinmetz and Brow, 2003). This indicates that the APT subcomplex is a specific snoRNA transcription termination complex. In addition, the APT subunit

Ref2p has been described as the recruiting factor for Glc7p to CPF (Nedea et al., 2003) and Ref2p is also a factor required for snoRNA transcription termination (Dheur et al., 2003). Therefore, we consider it likely that Glc7p as part of the subcomplex APT is involved in snoRNA transcription termination. It remains to be clarified if the APT subcomplex functions independently of the core-CPF in this process or whether the complete CPF complex is assembled on snoRNA terminators.

What could be the substrates for Glc7 in the process of snoRNA transcription termination?

The fact that Glc7p interacts with the APT subcomplex and is involved in the Nrd1-dependent snoRNA transcription termination pathway inspired us to test subunits of the APT and the Nrd1 complex as candidate substrates for the phosphatase Glc7p. Western analysis proved that neither Nrd1, Nab3, Sen1 nor Pti1p were substrates of Glc7p. In addition, Pta1p, the subunit of CPF that links the APT subcomplex to core-CPF, did not show an altered phosphorylation pattern in glc7 mutants either. This is remarkable, since Pta1p is a designated substrate of Glc7p in the process of polyadenylation (He and Moore, 2005), indicating that Glc7p must regulate the phosphorylation of a protein distinct from Pta1p in the process of snoRNA transcription termination.

Pcf11p, a subunit of CF IA, has been shown to be required in transcription termination of pre-mRNAs (Sadowski, 2003). Interestingly, it has a CTD interaction domain highly similar to the one of Nrd1p and was shown to bind preferentially to the phosphorylated CTD (Sadowski, 2003). The involvement of Pcf11p in snoRNA transcription termination was never tested, although two subunits of CF IA, Rna14p and Rna15p, have been reported to play a role in 3' end formation of snoRNAs (Morlando et al., 2002). The snoRNP core protein Nop1p interacts with all four CF IA subunits Pcf11p, Rna14p, Rna15p and Clp1p and the APT subunits Ssu72p, Ref2p, Pti1p and Swd2 (Morlando et al., 2004). In addition, Nop1p is required for 3' end formation of snoRNAs, suggesting that the interacting proteins for which a function in this process has not been demonstrated, like Pcf11 and Clp1p, might be involved (Morlando et al., 2004). Since we observed a strong GST-pulldown interaction of Glc7p with Pcf11p it would be interesting to test whether Pcf11p is involved in snoRNA transcription termination and is regulated by Glc7p in this process.

How transcription termination is triggered

Several cis- and trans-acting factors have been implicated in snoRNA and pre-mRNA transcription termination. However, the mechanism by which RNAP II is forced to dissociate from the DNA template and the nascent transcript remains elusive. Progress in understanding how RNAP II is triggered to release from the template has been made for transcription termination of pre-mRNAs. One quite mechanistical model implicates Pcf11p as a bridging factor between the nascent RNA and the CTD of RNAP II. Pcf11p dismantles RNAP II simply by restricting the distance RNAP II can transcribe further, away from the site where Pcf11p bound to the nascent RNA comparable to a dog on the leash. This ultimately causes RNAP II to dissociate from the template. As mentioned above Nrd1p and Pcf11p both have highly similar CTD interaction domains. Furthermore, Nrd1p, like Pcf11p is an RNA binding protein that binds to specific RNA motifs. It is tempting to speculate that Nrd1p might fill in the role of Pcf11p as a bridging factor in snoRNA transcription termination.

Transcription termination of pre-mRNAs further requires Rat1p, a 5'to 3' exonuclease that degrades the downstream cleavage product and is thought to torpedo RNAP II, facilitating its release (Kim et al., 2004b; Luo and Bentley, 2004). It is not known whether Rat1p is also required for snoRNA transcription termination.

3' end processing and transcription termination of pre-mRNAs are coupled. It has not been studied so far whether 3' end formation and transcription termination of snoRNAs is also coupled.

It is obvious that we still do not understand all aspects of transcription termination. It will be interesting to determine whether transcription termination of snoRNAs and of pre-mRNAs involve similar or even the same mechanisms to release RNAP II for new rounds of transcription.

4 Glc7p is involved in transcription elongation

4.1 Summary

During the transcription elongation phase RNAP II associates with a number of transcription elongation factors that regulate its processivity or remodel the chromatin structure. This actions are highly regulated to assure an efficient transcription of a gene. Here we report that the essential protein phosphatase Glc7p is required for poly(A)-dependent pausing. We show that several *glc7* mutant alleles are sensitive to the drug 6-azauracil (6AU), which is a phenotypic feature of transcription elongation defects. In addition, Glc7p genetically interacts with factors implicated in transcription elongation. These include Spt4p, a subunit of DSIF, and the PAF complex subunits Leo1p and Rtf1p. Based on these findings, we propose that Glc7p regulates certain aspects of transcription elongation. Interestingly, the snoRNA transcription termination defect we observe in *glc7* mutants is suppressed in *glc7/spt4*, *glc7/leo1* or *glc7/rtf1* double mutants. This could indicate a role for Spt4p, Leo1p and Rtf1p as antitermination factors in the process of snoRNA transcription and for Glc7p in coupling transcription elongation and termination of snoRNAs.

4.2 Introduction

Transcription of a gene by RNAP II can be divided into three main phases: initiation, elongation and termination (Saunders et al., 2006). The elongation phase of transcription by RNA polymerase II is highly regulated and tightly coupled to pre-mRNA processing (Maniatis and Reed, 2002; Proudfoot et al., 2002). The key component in this coupling is the C-terminal domain (CTD) of the largest subunit of RNAP II. Each transcription phase has a characteristic CTD phosphorylation pattern in which a specific subset of processing factors and also elongation factors is attracted to the CTD. Traveling along with the elongating RNAP II enables these factors to process the transcript as soon as it emerges from the transcription machinery. Vice versa, the status of processing can alter the transcriptional activity of RNAP II (Manley, 2002).

During elongation, RNAP II has to overcome nucleosome barriers or transient pausing caused by regulatory signals (Saunders et al., 2006). This is achieved with the help of positive and negative transcription elongation factors. Positive elongation

factors include TFIIS, TFIIF, FACT and the PAF complex (Belotserkovskaya and Reinberg, 2004; Conaway et al., 2003; Costa and Arndt, 2000; Mueller and Jaehning, 2002; Rondon et al., 2004; Tan et al., 1994). A transcription elongation factor that acts both positively and negatively is DSIF (DRB-sensitivity-inducing factor; (Lindstrom et al., 2003; Rondon et al., 2003). DSIF is composed of the proteins Spt4p and Spt5p (Hartzog et al., 1998; Wada et al., 1998a). Studies with Spt5p revealed that it is inhibitory to transcription only during early stages of transcription. During promoter-proximal pausing Spt5p and the CTD are phosphorylated by Ctk1p. Phosphorylated Spt5p then stimulates the transcription process by preventing premature termination and pausing (Bourgeois et al., 2002; Wada et al., 1998a). *In vitro* transcription assays and genetic studies revealed Spt4p as a positive transcription elongation factor (Rondon et al., 2003). The PAF complex consists of five subunits, Paf1p, Ctr9p, Cdc73p, Rtf1p and Leo1p (Krogan et al., 2002; Mueller and Jaehning, 2002; Squazzo et al., 2002). Originally, PAF was identified as a complex that associates with RNAP II (Shi et al., 1996). PAF is crucial in regulating monoubiquitylation and methylation of several histones (Wood et al., 2003) thus supports transcription elongation as a factor that remodels chromatin.

Sensitivity to 6-azauracil is a phenotypic landmark of transcription elongation defects. The drug 6AU reduces intracellular levels of UTP and GTP thus depriving RNAP II of substrates for transcription. Mutations in many elongation factors cause hypersensitivity to 6AU, because the lack of their positive influence on transcription processivity renders RNAP II more vulnerable to nucleotide reduction. Subunits of the DSIF and the PAF complex show varying degrees of sensitivity to 6 AU (Hartzog et al., 1998; Squazzo et al., 2002).

Different processing events can influence each other. For example, termination of pre-mRNA transcription is coupled to 3' end formation (Proudfoot, 2000a). Several recent reports present data that indicate a coupling of transcription elongation to transcription termination (Bucheli and Buratowski, 2005; Kaplan et al., 2005). It was found that Spt4p and some other elongation factors have properties of antitermination factors at the GAL10 – GAL7 locus (Kaplan et al., 2005). Kaplan and co-workers suggested that a reorganization of elongation factors associated with RNAP II past the polyadenylation signal would facilitate transcription termination (Kaplan et al., 2005). It was observed that the PAF complex crosslinks to genes throughout the transcribed regions of genes up to the polyadenylation site (Buratowski, 2005). While the

polymerase and several other elongation factors continue to transcribe, very little PAF cross-linking is seen 3' to the poly(A) site (Ahn et al., 2004; Kim et al., 2004a). If the PAF complex acts as a positive elongation factor, polymerases downstream of the polyadenylation site that are no longer associated with PAF could become more competent for termination. This again suggests that a reorganization of the RNAP II transcription complex is required to allow transcription termination.

Accumulating evidence suggests that transcription elongation factors are also required for 3'end formation of snoRNAs, a class of non-coding RNAP II transcripts. First, Sheldon and co-workers presented evidence that defects in the Paf1 complex lead to the accumulation of readthrough products at snoRNA loci (Sheldon et al., 2005). Second, Spt5p copurifies with the Nrd1 complex, a complex required for snoRNA transcription termination, suggesting tight coupling of elongation and termination of snoRNAs (Vasiljeva and Buratowski, 2006).

In this work we analysed the function of Glc7p during transcription elongation and termination of pre-mRNAs and snoRNAs. We show that Glc7p is essential for snoRNA transcription termination. In addition our results show that Glc7 interacts genetically with the elongation factors Spt4p, Rtf1p and Leo1p and that the transcription termination defect of *glc7* mutants can be suppressed by deletion of Spt4p, Rtf1p and Leo1p suggesting an antiterminator function of these elongation factors in the process of snoRNA termination.

4.3 Results

4.3.1 Poly(A)-dependent pausing of RNAP II is reduced in glc7-12 mutants

In mammals, the poly(A) signal drives transcription termination in two steps: poly(A)-dependent pausing of RNAP II followed by its poly(A)-dependent release. The CTD of RNAP II is required for poly(A)-dependent termination, but the body of the polymerase is sufficient for poly(A)-dependent pausing in the absence of the CTD (Park et al., 2004). While analyzing the results of our CYC1 TRO experiments we realized that WT cells consistently showed a very strong signal over probe P2 at 37ºC (Fig. 3.4 c). Probe P2 covers the poly(A) site of CYC1, indicating that a large amount of RNAP II molecules must have been stalled at this position. This enhanced pausing at the poly(A) signal seemed to be induced by the heat shock because at room temperature (RT) only moderate pausing is observed in the WT. Interestingly, pausing at the poly(A) site in *glc7-12* mutants is reduced dramatically compared to the WT, indicating that RNAP II more often reads through the pause site without stalling. However, reduced stalling at the poly(A) site was not sufficient to interfere with transcription termination in *glc7-12* mutants (Fig. 3.4).

The poly(A) site is not the only transcriptional pause site during transcription of a gene. Another well known pause site is located at the 5' end of genes at the transition of transcription initiation to transcription elongation. Transcription elongation factors modulate the processivity of RNAP II both negatively and positively at this pause site. This raised the question whether Glc7p might regulate transcription elongation factors to adapt RNAP II to cellular signals and is therefore involved in transcription elongation. Another possibility to explain the lack of stalling over the poly(A) site in *glc7-12* mutants is that Glc7p directly modifies the body of RNAP II by dephosphorylation, since RNAP II without the CTD is sufficient for pausing.

It is intriguing that *ssu72-2* mutant cells also showed an abnormal density of RNAP II molecules at the poly(A) site (Dichtl et al., 2002a). In contrast to *glc7-12* mutant cells however, ssu72-2 mutants display enhanced pausing of RNAP II at this site compared to the WT. Further investigation of this effect has lead to an implication of Ssu72p in transcription elongation. In addition, *ssu72-2* mutants have a defect in snoRNA transcription termination (Dichtl et al., 2002a). Therefore it is likely that Ssu72 and Glc7p share similar functions in these processes.

4.3.2 Growth of glc7 mutant strains is affected by 6-azauracil

To investigate whether Glc7p could be involved in regulating transcription elongation we applied a commonly used assay to test for growth of *glc7* mutants on the drug 6-azauracil (6AU). 6AU inhibits both the orotidylic acid decarboxylase and the IMP dehydrogenase (IMPDH), rate-limiting enzymes in the UTP and GTP biosynthetic pathways (Hampsey, 1997). 6AU diminishes the intracellular pools of UTP and GTP which is thought to affect the *in vivo* transcription elongation rate by depleting RNAP II for these nucleoside triphosphates. This is believed to reflect the need for an optimally functioning transcription machinery under conditions of lowered intracellular GTP and UTP levels. Whereas WT cells can tolerate 6AU up to a certain concentration, cells with mutations in transcription elongation factors react to the drug and either decrease or accelerate growth. This is because transcription elongation factors influence the transcription rate and the processivity of RNAP II either positively or negatively depending on the factor. For this reason, sensitivity of cells to 6AU has been correlated with mutations in transcription elongation factors which influence transcription positively, whereas resistance to 6AU implies a repressive function of the respective transcription factor (Costa and Arndt, 2000; Nakanishi et al., 1995; Powell and Reines, 1996; Squazzo et al., 2002).

We tested the WT, *glc7-10*, *glc7-12*, *glc7-13* and *glc7-5* strains for growth on medium containing different concentrations of 6AU (0, 10, 25, 50 and 100µg/ml) to see how strongly the mutants react to the drug (Fig. 4.1). In addition to growth at RT, cells were also incubated at semi-permissive and non-permissive temperatures. Because uracil in the growth medium can rescue the growth defects caused by 6AU, we grew the cells on synthetic medium lacking uracil. Growth defects become readily visible at higher dilutions, thus serial dilutions of cells spotted with highest concentrations on the left hand side of each column are depicted.

Significant effects of 6AU on growth of all the mutants tested compared to the WT could be seen even at RT. *glc7-10* and *glc7-12* are resistant to 6AU, which becomes most apparent at high concentrations of the drug (50 and 100 µg/ml). In contrast, *glc7-13* and *glc7-5* are sensitive to 6AU, which is evident at 6AU concentrations of 25 µg/ml and higher. These effects become even more obvious at 30º C, the semi-permissive temperature for *glc7-12*, *glc7-13* and *glc7-5* and at 33º C, the semi-permissive temperature for *glc7-10*. Because 33º C is the restrictive temperature for

glc7-12, *glc7-13* and *glc7-5*, only *glc7-10* was analyzed at higher temperatures. At 37º C, 6AU even rescues growth of *glc7-10* which is unable to grow at this temperature without 6AU in this range of dilutions (a 10 fold higher concentration of cells is able to grow at this temperature, results not shown).

In conclusion, we observe very strong effects of 6AU on the growth of all *glc7* mutants tested, suggesting an involvement of Glc7p in the regulation of transcription elongation. Our observation that distinct alleles are either sensitive or resistant to 6AU and therefore most likely influence transcription by RNAP II positively or negatively, lets us propose that Glc7p might regulate transcription elongation at several different steps.

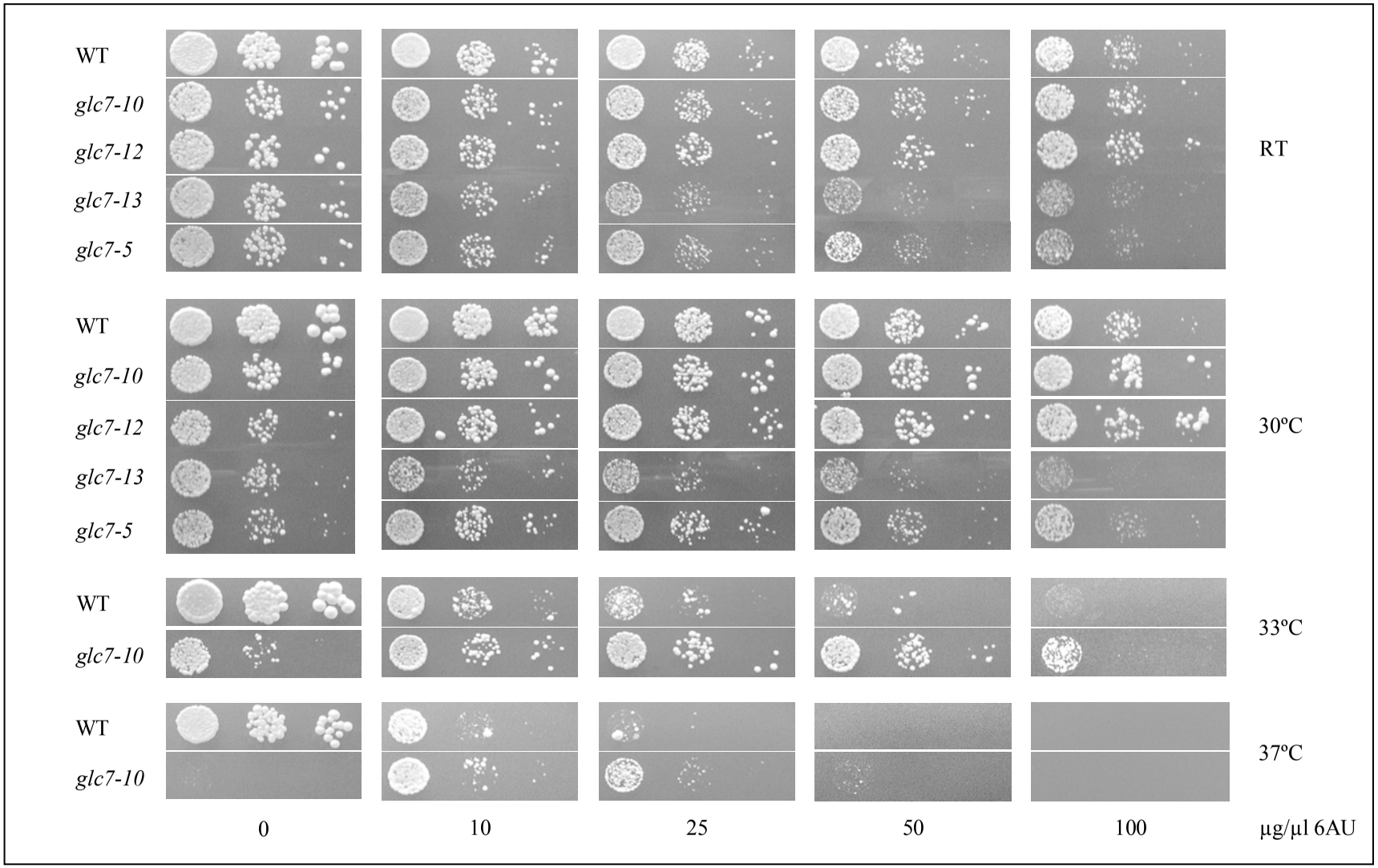


Figure 4.1. Influence of 6AU on growth of *glc7* mutant strains.
WT, *glc7-10*, *glc7-12*, *glc7-13* and *glc7-5* strains were transformed with pRS416 (URA3-CEN). Ten-fold serial dilutions of these strains grown at RT in synthetic medium (SD-ura; 0,67% yeast nitrogen base, 2% glucose, 1x complete drop out mix lacking uracil) were spotted on the same medium (plus 2% agar) containing the indicated amounts of 6AU. The time of growth was 3 days for all panels at the respective temperatures indicated on the right.

4.3.3 Transcription of Pur5 is not altered by 6AU in glc7 mutant strains

To further confirm that Glc7p is involved in transcription elongation as suggested by the growth effects caused by 6AU, we decided to test the induction of PUR5. In response to the application of IMPDH inhibitors such as 6AU, WT yeast strains induce transcription of PUR5, one of four genes encoding IMPDH-related enzymes (Shaw and Reines, 2000). Mutations that cause transcription elongation defects and 6AU sensitivity also prevent the induction of PUR5 transcription in response to 6AU treatment. Thus, a 6AU-sensitive phenotype coupled with the inability to induce PUR5 in response to 6AU is suggestive of a defect in transcription elongation (Shaw and Reines, 2000).

We investigated the induction of PUR5 before and after 6AU treatment by Northern analysis. To this end, we collected total RNA from WT, *glc7-10*, *glc7-12*, *glc7-13* and *glc7-5* after one or two hours of 6AU treatment. As a control, total RNA was also collected without 6AU treatment.

As expected, WT cells showed a strong induction of PUR5 after treatment with 6AU (Howard et al., 2003; Shaw and Reines, 2000; Squazzo et al., 2002; Fig. 4.2). In contrast, induction of PUR5 was greatly attenuated in *glc7-10* and *glc7-12* cells and also to a lesser degree in *glc7-13*, indicating that the elongation machinery was impaired. This result, together with the 6AU phenotype, strongly suggests a function for Glc7p in transcription elongation.

One of the *glc7* mutants, *glc7-5*, was able to induce WT levels of PUR5 although growth of *glc7-5* is very sensitive to 6AU. Interestingly, it was found that mutations in Spt4 and Spt5, which are well known transcription elongation factors, were able to induce PUR5 to levels that approached those seen in WT cells, although these mutants were sensitive to 6AU (Howard et al., 2003). This suggested the existence of two general classes of drug-sensitive elongation mutants: those that are defective in PUR5 induction and those that are not. In conclusion, *glc7-5* might belong to the latter class and therefore does not challenge the observations with the other *glc7* mutants.

Altogether, our results support the idea that Glc7p is involved in transcription elongation. Furthermore, these data suggest that not only mutants sensitive to 6AU exhibit reduced induction of PUR5, but also mutants that are insensitive to 6AU. The reason for this remains to be uncovered.

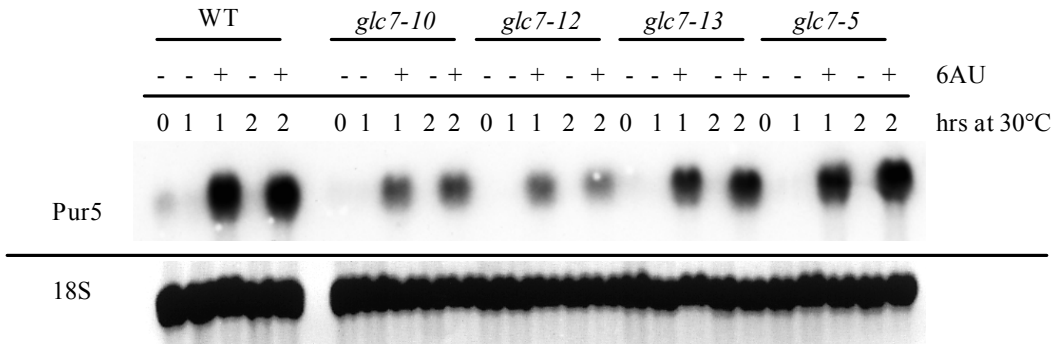


Figure 4.2. Induction of PUR5 is affected in *glc7* mutant strains
Northern analysis of 20µg total RNA extracted from WT, *glc7-10*, *glc7-12*, *glc7-13* and *glc7-5* cells following growth in synthetic medium (SD-ura; 0,67% yeast nitrogen base, 2% glucose, 1x complete drop out mix lacking uracil) with (+) or without (-) addition of 6AU (50µg/ml) at 30º C for the indicated times. 18S serves as loading control.

4.3.4 Glc7p genetically interacts with transcription elongation factors

To better understand the role of Glc7p in the process of transcription elongation, we decided to look for functional interactions of Glc7p with transcription elongation factors *in vivo*. For this, we generated a set of double mutants containing a *glc7* mutation and a deletion of either of the transcription elongation factors Spt4p, Rtf1p or Leo1p.

4.3.4.1 Genetic interactions with Spt4p

Spt4p is found together with Spt5p in a single, highly conserved complex, the Spt4/5 complex (Hartzog et al., 1998). The human homologue of the Spt4/5 complex is DSIF (Wada et al., 1998a). One of the most interesting features of DSIF is that it appears to act as both a negative and positive regulator of RNAP II elongation (Renner et al., 2001; Rondon et al., 2003; Yamaguchi et al., 1999). Since we found two different growth responses to 6AU when testing distinct *glc7* mutant alleles, we suspected that Glc7p also can regulate transcription elongation both positively and negatively. In addition, *in vitro* studies indicate that DSIF increases the pausing of RNAP II (Renner et al., 2001; Yamaguchi et al., 1999). Our observations that *glc7-12* decreases pausing

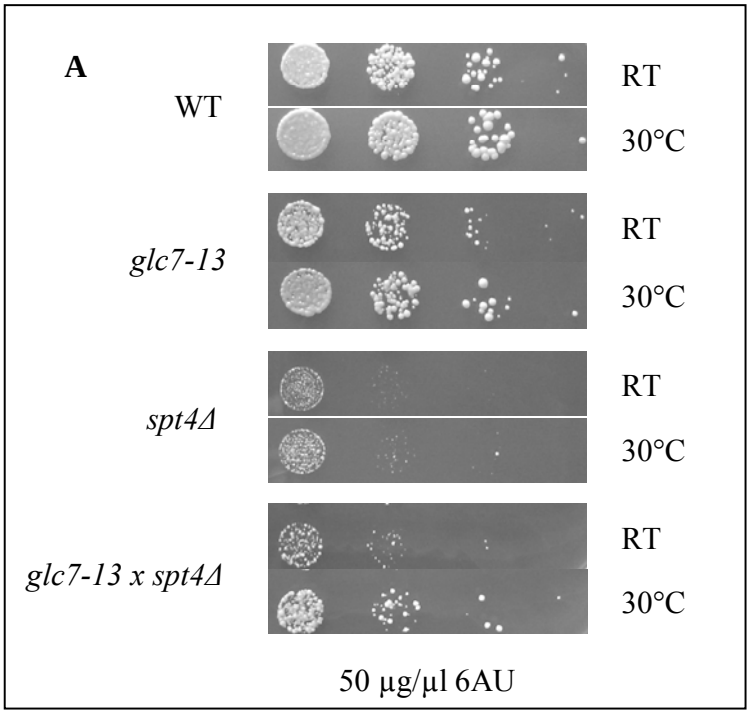
at the poly(A) site suggest that Glc7p could function similar to DSIF in this respect. Thus, the Spt4/5 complex seemed a likely interaction partner for Glc7p.

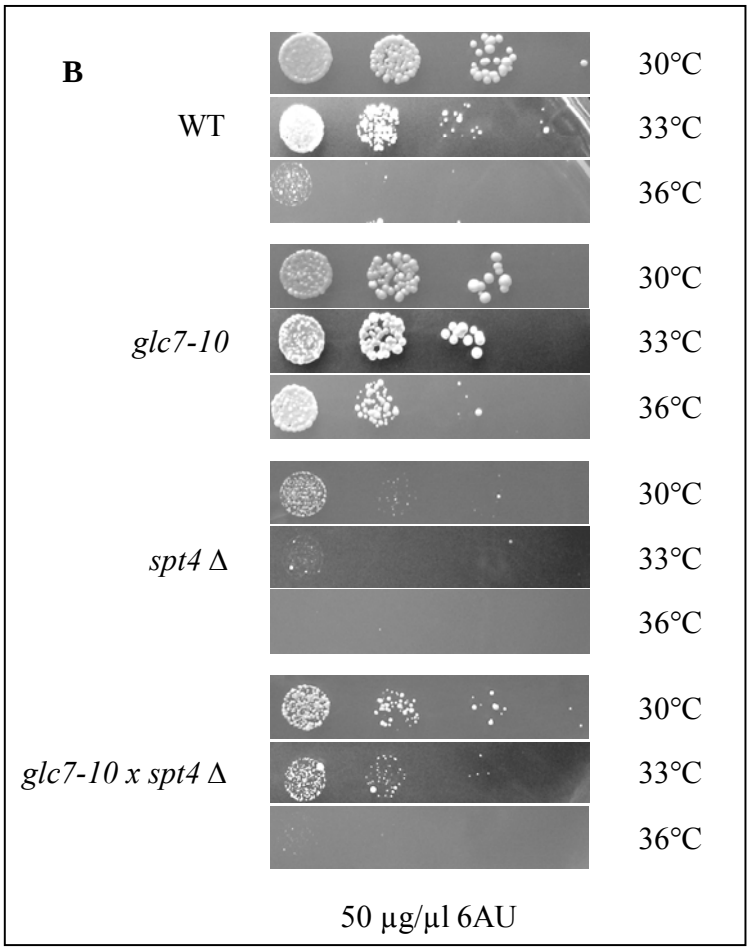
We found that both strains were sensitive to 6AU, however to varying degrees. *spt4Δ* was very sensitive to 6AU even at RT, whereas *glc7-13* only showed a slight sensitivity to 6AU in this experiment. When we crossed a spt4 deletion strain with the *glc7-13* mutant strain, the strong sensitivity of *spt4Δ* to 6AU could be partially rescued in the double mutant *spt4 Δ /glc7-13* at RT and 30º C as indicated by the better growth of the double mutant compared to the single *spt4Δ* mutant (Fig. 4.4 a). This result suggests an interaction between these genes. However, growth was not rescued at higher temperatures, indicating that rescuing defects in transcription elongation might not be enough to restore growth since Glc7p is involved in many other processes.

In the presence of 6AU growth of *glc7-10* was enhanced compared to the WT. The *glc7-10* allele was even able to suppress the 6AU sensitivity of *spt4Δ* at RT, 30º C and 33° C in the *glc7-10/spt4Δ* double mutant (Fig. 4.4 b). However, suppression of the *spt4Δ* growth defect on 6AU was more reduced when the temperature conditions for *glc7-10* became more stringent. Complete loss of growth was observed at 36º C, a temperature that still allowed growth of *glc7-10*. Therefore, deletion of spt4 in *glc7-10* strains is lethal under conditions that normally support growth of *glc7-10*, again indicating genetic interactions between Spt4p and Glc7p.

Taken together, we observed opposite effects concerning growth in *spt4Δ/glc7* double mutants, depending on the growth behavior of the single *glc7* mutant on 6AU: the 6AU sensitive *glc7-13* allele improved growth of the double mutant, whereas the 6AU resistant glc-10 allele did not. In other words, *glc7-13* partially suppressed the growth defect of *spt4*, whereas *spt4* suppressed growth of *glc7-10*.

Therefore, we conclude that Glc7p genetically interacts with the Spt4/5 complex. These observations strongly suggest an *in vivo* role for Glc7p in transcription elongation.


A
WT
RT
30°C
glc7-13
RT
30°C
spt4Δ
RT
30°C
glc7-13 x spt4Δ
RT
30°C
50 µg/µl 6AU


B
WT
30°C
33°C
36°C
glc7-10
30°C
33°C
36°C
spt4 Δ
30°C
33°C
36°C
glc7-10 x spt4 Δ
30°C
33°C
36°C
50 µg/µl 6AU

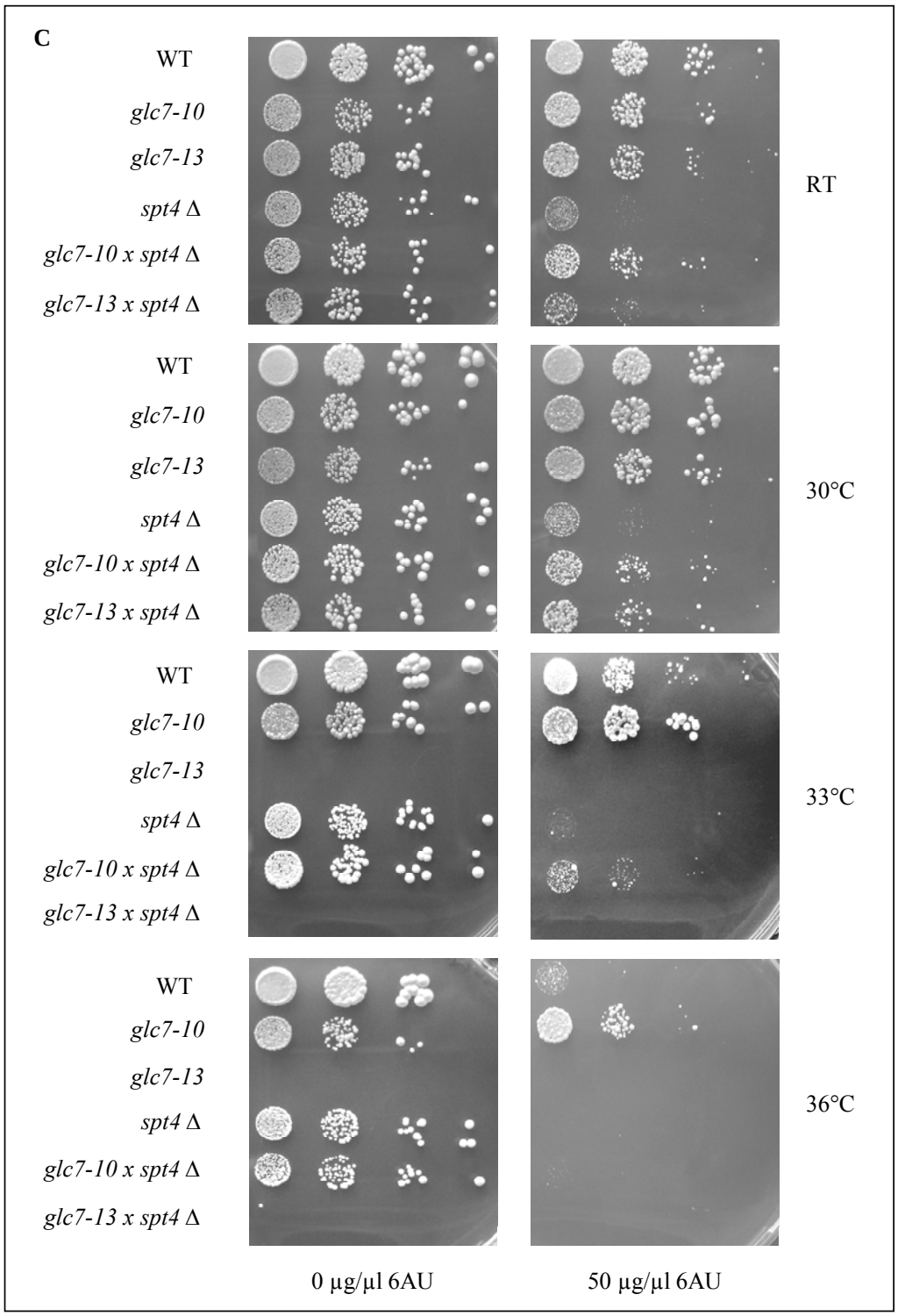


Figure 4.3. Genetic interactions of Glc7 with Spt4 revealed by growth on 6AU.
WT, *glc7-10*, *glc7-13* and *spt4Δ* strains were transformed with pRS416 (URA3-CEN). Ten-fold serial dilutions of these strains grown at RT in synthetic medium (SD-ura; 0,67% yeast nitrogen base, 2% glucose, 1x complete drop out mix lacking uracil) were spotted on the same medium (plus 2% agar) with or without 6AU. The time of growth was 3 days for all panels at the respective temperatures indicated on the right. A) Summary of interactions between glc7-13 and spt4Δ. B) Summary of interactions between glc7-10 and spt4Δ. C) Complete experiment showing growth of strains indicated on the left also on synthetic medium without 6AU.

4.3.4.2 The phosphorylation state of Spt5 is not regulated by Glc7p

Spt5p is a phosphoprotein (Ivanov et al., 2000). Regulation of the phosphorylation state of Spt5p is important for progression of the transcription cycle. It was shown that DSIF binds to RNAP II and effectively blocks transcription elongation in early stages of transcription (Wada et al., 1998a). This block is relieved by P-TEFb, the human homologue of the yeast CTD kinase Ctk1p. P-TEFb phosphorylates both the CTD and Spt5 (Ivanov et al., 2000; Kim and Sharp, 2001; Wada et al., 1998b). Following the pTEFb-mediated phosphorylation DSIF remains associated with the elongating polymerase and can apparently stimulate the elongation process (Ping and Rana, 2001).

The genetic interactions we found between Glc7p and Spt4p could indicate that Glc7p might interfere with transcription elongation by regulating the phosphorylation state of Spt5p. To this end, we generated constructs with a ProteinA-tag at the C-terminus of Spt5p in WT, *glc7-12* and *glc7-13* mutant strains. This allowed us to affinity-purify Spt5p from these strains and subsequently investigate phosphorylation of Spt5p by Western analysis with antibodies directed against phospho-serine and threonine. The results showed no difference in the phosphorylation state of Spt5p between WT and the *glc7* mutant strains after shift to the non-permissive temperature (B. Paguet, data not shown) indicating that Glc7p does not dephosphorylate Spt5p.

In conclusion, Spt5 does not seem to be the target of Glc7p in transcription elongation and therefore Glc7p does not regulate pausing of RNAP II early in transcription. Nevertheless, the genetic interactions of Glc7p with Spt4 suggest that Glc7p could influence transcription elongation through elongation factors that associate with the Spt4/5 complex.

4.3.4.3 Genetic interactions with subunits of the PAF1 complex

Previous studies have shown that the Spt4-Spt5 complex co-immunopurifies and genetically interacts with the Paf1 complex (Mueller and Jaehning, 2002; Squazzo et al., 2002). The Paf1 complex consists of the subunits Paf1p, Ctr9p, Cdc73p, Rtf1p and Leo1p (Betz et al., 2002; Squazzo et al., 2002). It has been shown that mutations in Paf1p and Leo1p can suppress a cold-sensitive spt5 mutation (Squazzo et al., 2002) and a deletion of Rtf1p showed a range of synthetic phenotypes when combined with mutations in spt4 and spt5 (Costa and Arndt, 2000). In addition, members of the Paf

complex exhibit 6AU sensitivity and diminished PUR5 induction (Squazzo et al., 2002). Finally, ChiP experiments demonstrated that the Paf complex is associated with RNAP II at promoters and coding regions of transcriptionally active genes (Kim et al., 2004a; Penheiter et al., 2005). For all of these reasons, the Paf complex has been implicated in transcription elongation.

To find out whether Glc7p also genetically interacts with the Paf1 complex, we crossed the *glc7-10*, *glc7-12*, *glc7-13* and *glc7-5* mutant strains to deletion mutants of Paf1p, Ctr9p, Leo1p and Rtf1p in all combinations. This approach led to the recovery of three viable double mutants: *glc7-12/leo1Δ, glc7-10/rtf1Δ* and *glc7-13/rtf1Δ.* We verified that these were indeed double mutants by Southern analysis and PCR analysis of the genomic regions.

When testing the *glc7-12/leoΔ* double mutant for growth at various temperatures, we observed that temperature sensitivity of glc7-12 at 33º C is slightly suppressed by *leo1Δ,* indicating genetic interactions between these genes (Fig. 4.4 a). This suppression is strongly enhanced when *glc7-12/leo1Δ* is grown on medium containing 6AU, demonstrating that this genetic interaction is important for the process of transcription elongation. Interestingly, the *glc7-12/leo1Δ* double mutant is not viable at higher temperatures. This might be due to the fact that Glc7p is involved in many cellular processes which are not rescued by this interaction. Consequently, *leo1Δ* is able to partially suppress the temperature sensitivity of the *glc7-12* mutation. Nevertheless, we clearly see genetic interactions between *glc7-12* and *leo1Δ* that are amplified by 6AU, strongly suggesting that Glc7p regulates an aspect of transcription elongation.

An even more remarkable suppression of temperature-sensitivity is observed in the *glc7-13/rtf1Δ* double mutant. *rtf1Δ* fully suppresses lethality of *glc7-13* at 33° C independently of 6AU, because the *glc7-13/rtf1Δ* double mutant grows equally well on medium with or without 6AU (Fig. 4.4 b). Again, suppression of temperature-sensitivity is limited to 33º C and *glc7-13/rtf1Δ* is not able to grow at higher temperatures. Therefore, we think that Rtf1p genetically interacts with Glc7p as well. Considering the 6AU independence of the temperature-sensitivity suppression this interaction might indicate an involvement of these factors not only in transcription elongation but in additional processes.

In contrast to the *glc7-13/rtf1Δ* and *glc7-12/leo1Δ* double mutant, the *glc7-10/rtfΔ* double mutant is not able to grow at the restrictive temperature for *glc7-10* in the

presence of 6AU (Fig. 4.4 c). The single mutants *rtf1Δ* and *glc7-10* show varying degrees of 6AU resistance at this temperature compared to the WT. However, these effects are not sufficient to rescue growth under these conditions, whereas on medium without 6AU the *glc7-10/rtfΔ* double mutant is viable.

In summary, the described double mutants reveal genetic interactions between different *glc7* mutant alleles and the transcription elongation factors Rtf1p and Leo1p and imply a regulatory function for Glc7p in the process of transcription elongation.

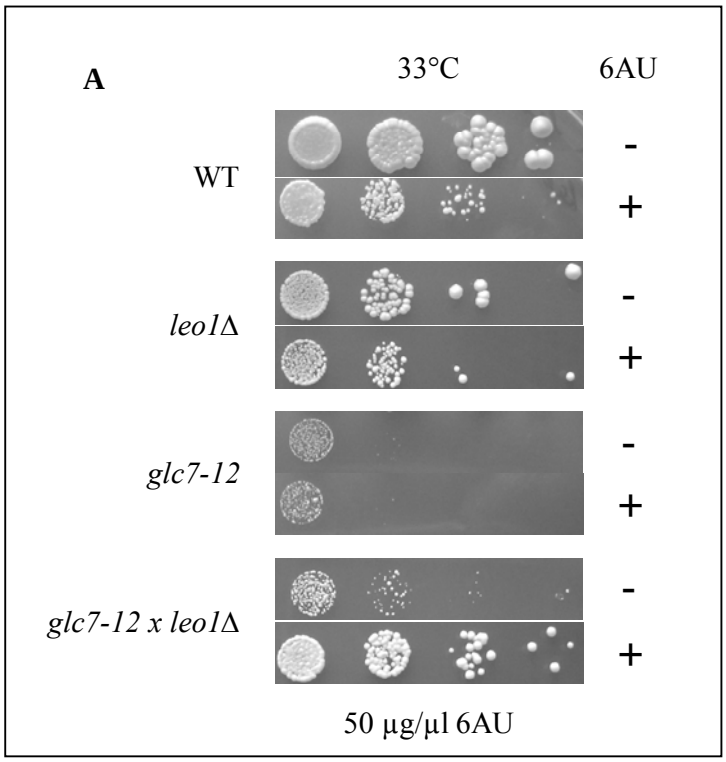
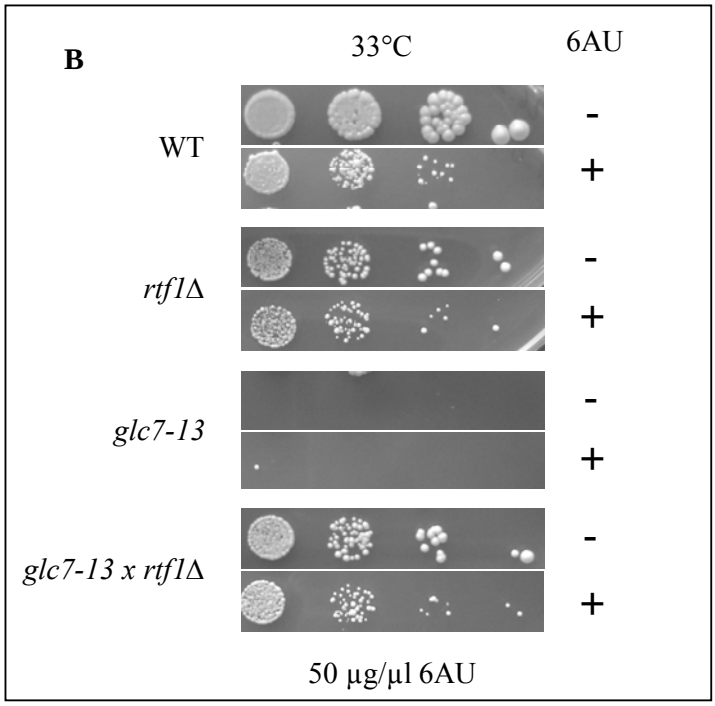
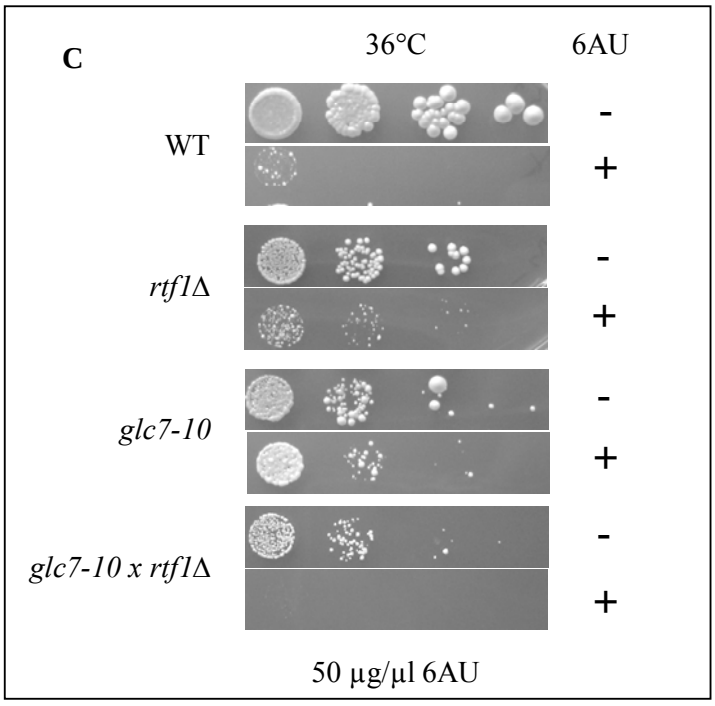
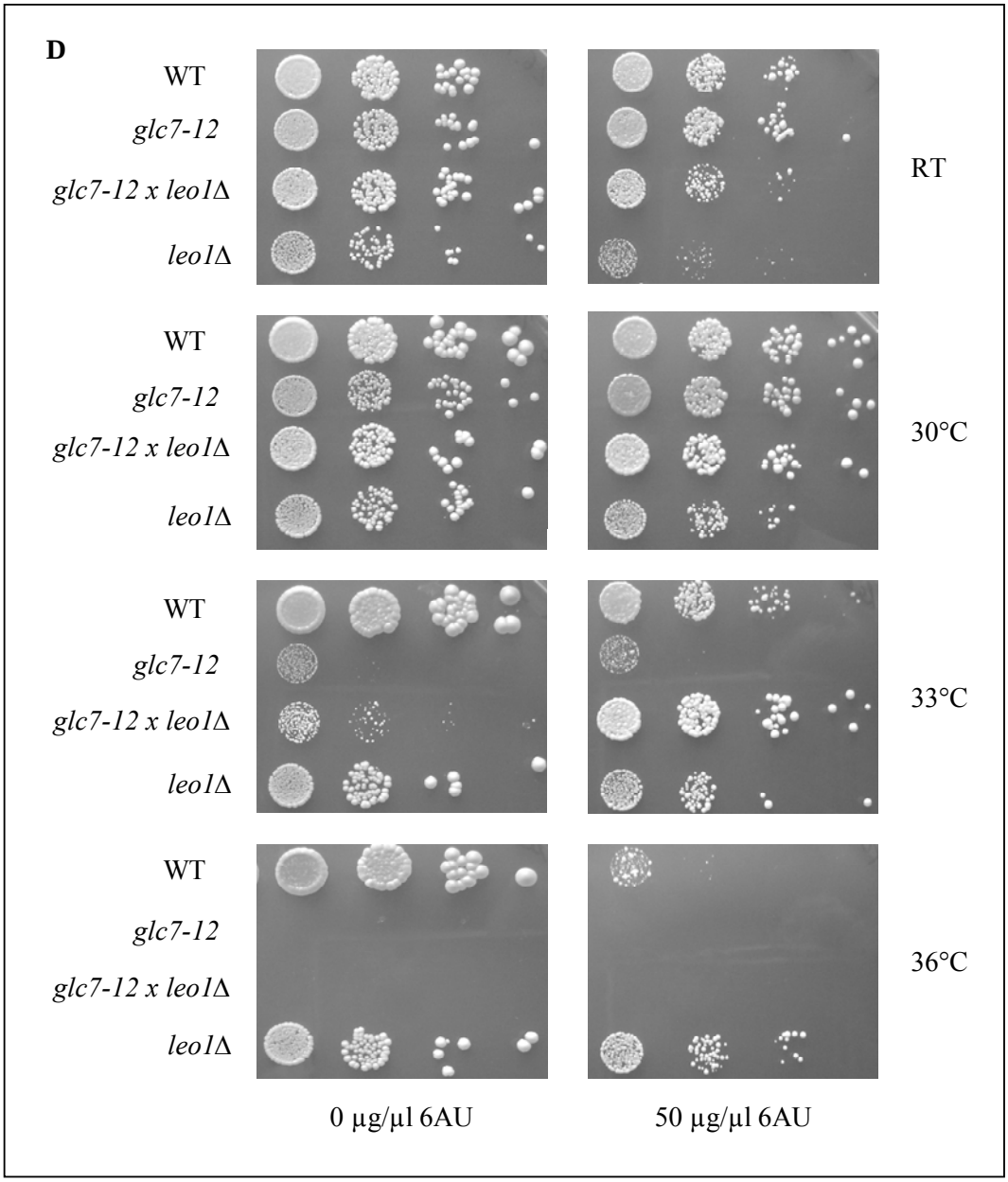


Figure 4.4.

For detailed information see page 104


B
33°C
6AU
WT
rtf1Δ
glc7-13
glc7-13 x rtf1Δ
-
+
50 µg/µl 6AU


C
36°C
6AU
WT
rtf1Δ
glc7-10
glc7-10 x rtf1Δ
-
+
50 µg/µl 6AU


D
WT
glc7-12
glc7-12 x leo1Δ
leo1Δ
RT
30°C
33°C
36°C
0 µg/µl 6AU
50 µg/µl 6AU

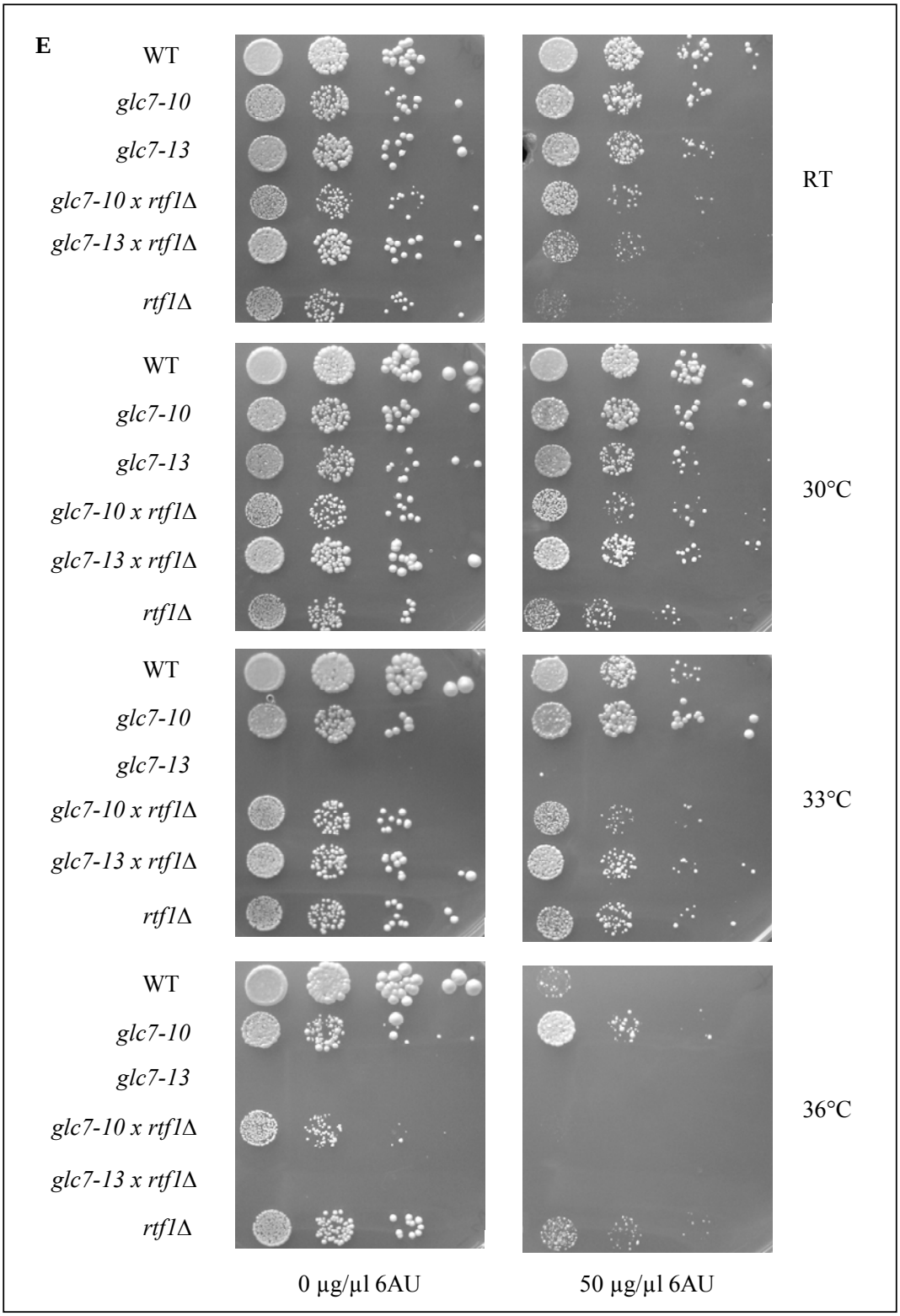


Figure 4.4. Genetic interactions of Glc7 with the Paf complex subunits Leo1 and Rtf1.
WT, *glc7-10*, *glc7-12, glc7-13*, *leo1Δ* and *rtf1Δ* strains were transformed with pRS416 (URA3-CEN). Ten-fold serial dilutions of these strains grown at RT in synthetic medium (SD-ura; 0,67% yeast nitrogen base, 2% glucose, 1x complete drop out mix lacking uracil) were spotted on the same medium (plus 2% agar) with or without 6AU. The time of growth was 3 days for all panels at the respective

temperatures indicated on the right. A) Summary of interactions between *glc7-12* and *leo1Δ*. B) Summary of interactions between *glc7-13* and *rtf1Δ*. C) Summary of interactions between *glc7-10* and *rtf1Δ* D) Complete experiment showing growth of strains *glc7-12*, *leo1Δ* and the double mutant strain *glc7-12 x leo1Δ* as indicated on the left for different temperatures as indicated on the right. E) Complete experiment showing growth of strains *glc7-10*, *glc7-13*, *rtf1Δ* and the double mutant strains *glc7-10 x rtf1Δ* and *glc7-13 x rtf1Δ* as indicated on the left for different temperatures as indicated on the right.

4.3.5 Glc7 mutants are not inositol auxotroph

Inositol auxotrophy is often indicative of defects in the transcription apparatus, presumably because the INO1 gene is extremely sensitive to transcriptional perturbations in general. An Ino$^-$ phenotype is often an important clue that a mutant is defective in a component of the RNAP II transcriptional machinery (Hampsey, 1997). Interestingly, the *paf1Δ* and *ctr9Δ* mutations exhibit strong Ino$^-$ phenotypes (Squazzo et al., 2002), indicating that the Paf1 complex is part of the RNAP II transcriptional machinery. However mutations in other members of the Paf1 complex, such as *leo1Δ* and *rtf1Δ* show only weak inositol auxotrophy.

We therefore asked whether mutations in Glc7p have an Ino$^-$ phenotype, linking Glc7p more closely to the RNAP II transcriptional machinery and also to the Paf1 complex.

Inositol auxotrophy is scored on synthetic medium lacking inositol. It is important to establish that impaired growth in the absence of inositol is indeed a consequence of inositol limitation; this is done by scoring for the ability of exogenous inositol to rescue the Ino$^-$ phenotype (Hampsey, 1997). None of the *glc7* mutant strains that we tested were auxotroph for inositol (Fig. 4.5), suggesting that Glc7p is probably not permanently attached to the RNAP II general transcriptional machinery. As a control we used the *ctr9Δ* mutant strain, which shows strong inositol auxotrophy. Interestingly, *glc7-12* and *glc7-5* mutant strains grew better on medium lacking inositol. The cause for this was not known. Subsequently, microarray analysis with the *glc7-12* mutant strain revealed that expression of INO1 was upregulated. The reason for this upregulation is unclear, but a higher level of INO1 could account for the enhanced growth of at least the *glc7-12* mutant on medium lacking inositol.

Altogether, our genetic tests suggest that Glc7p might only transiently interact with the RNAP II transcriptional machinery where it interacts and regulates the Paf1 complex.

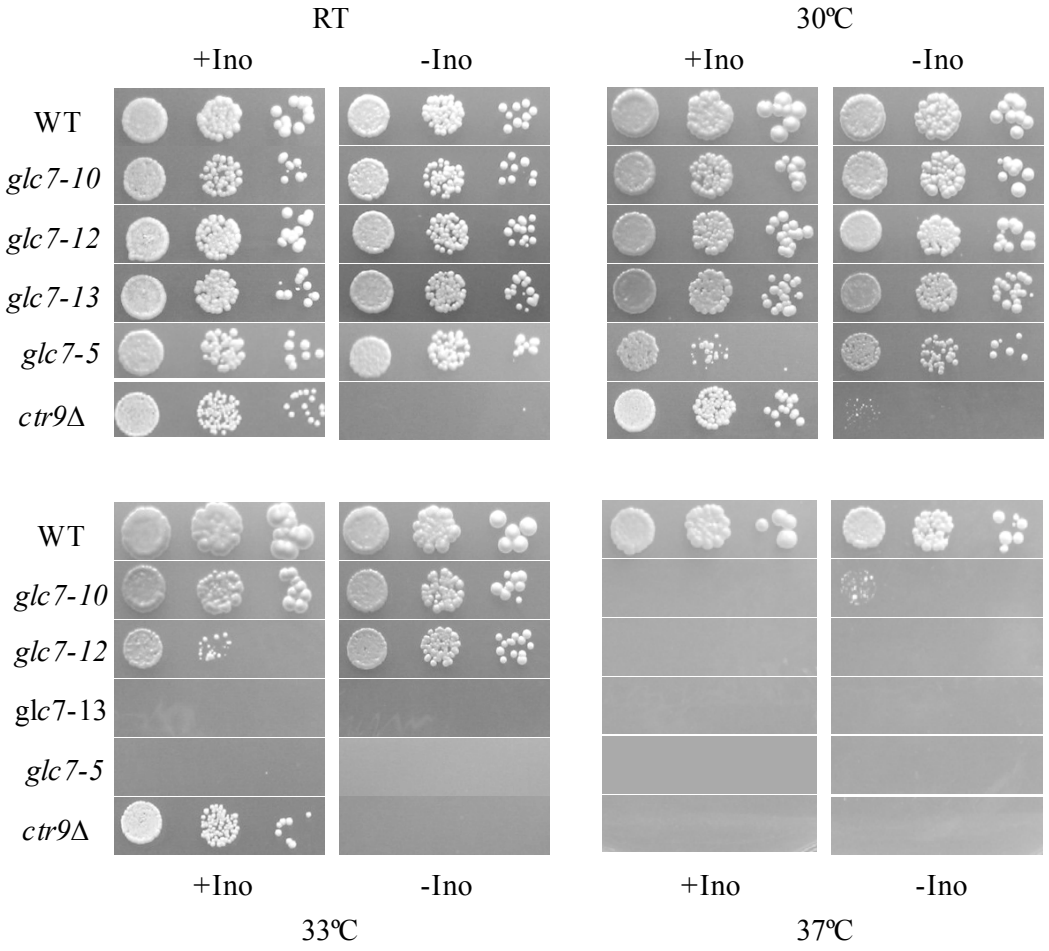


Figure 4.5. Growth of *glc7* mutants is not affected on medium lacking inositol.
Ten-fold serial dilutions of the strains indicated on the left were grown at RT in YPD and spotted on synthetic medium without inositol or on control medium containing 10mg/l inositol. The time of growth was 3 days for all panels at RT, 30º C, 33º C or 37º C as indicated. *ctr9Δ* was used as a positive control.

4.3.6 The snoRNA transcription termination defect of glc7 mutant cells is suppressed by transcription elongation factors

Recently, there have been several reports indicating that transcription elongation factors can influence transcription termination (Kaplan et al., 2005; Penheiter et al., 2005; Sheldon et al., 2005). First, an *spt4Δ* mutant strain showed reduced levels of read-through at the poly(A) site of Gal 10, indicating that Spt4 may act as an anti-terminator (Kaplan et al., 2005). In addition, ChIP experiments have demonstrated that Paf1C is required for the efficient recruitment of Nrd1 to a snoRNA gene and that

deletion of all subunits of the Paf1 complex leads to the accumulation of 3' extended snoRNA transcripts (Sheldon et al., 2005). Finally, it has been shown that there is an exchange of elongation factors for 3' end processing factors at the 3' end of genes and therefore a reorganization of the RNAP II holoenzyme in response to 3' end formation signals during transcription termination has been proposed (Kaplan et al., 2005; Kim et al., 2004a).

Considering these data we asked whether the partial suppression of the *ts* phenotype of *glc7-12* and *glc7-13* mutant strains in the *glc7-13/spt4Δ, glc7-13/rtf1Δ* and *glc7-12/leo1Δ* double mutants was due to the fact that deletion of these elongation factors would influence snoRNA transcription termination in the double mutant.

To study snoRNA transcription termination in the double mutants we performed Northern analysis with the same probes that led to the finding of 3' extended snoRNAs in the *glc7* mutants (see Chapter 3). To this end, we extracted total RNA before and after shift to 33ºC from the double mutant strains and also from the WT and respective single mutant strains. We shifted the cells to 33º C in this experiment and not to 37º C because we found that suppression of the *ts* phenotype of *glc7-12* and *glc7-13* mutant strains was limited to 33º C. As shown in Fig. 4.6 a, read-through products are readily detected for *glc7-12* and *glc7-13* mutants also at 33º C, not only at 37º C. In *glc7-10* mutants read-through products are also visible at 33º C, however to a very low degree. In contrast, we could not detect any read-through products for any of the elongation factor mutants *spt4*, *leo1* or *rtf1* in the Northern analysis with several probes. We also did not observe read-through products for *spt4*, *leo1* or *rtf1* when cells were grown at 37ºC (results not shown).

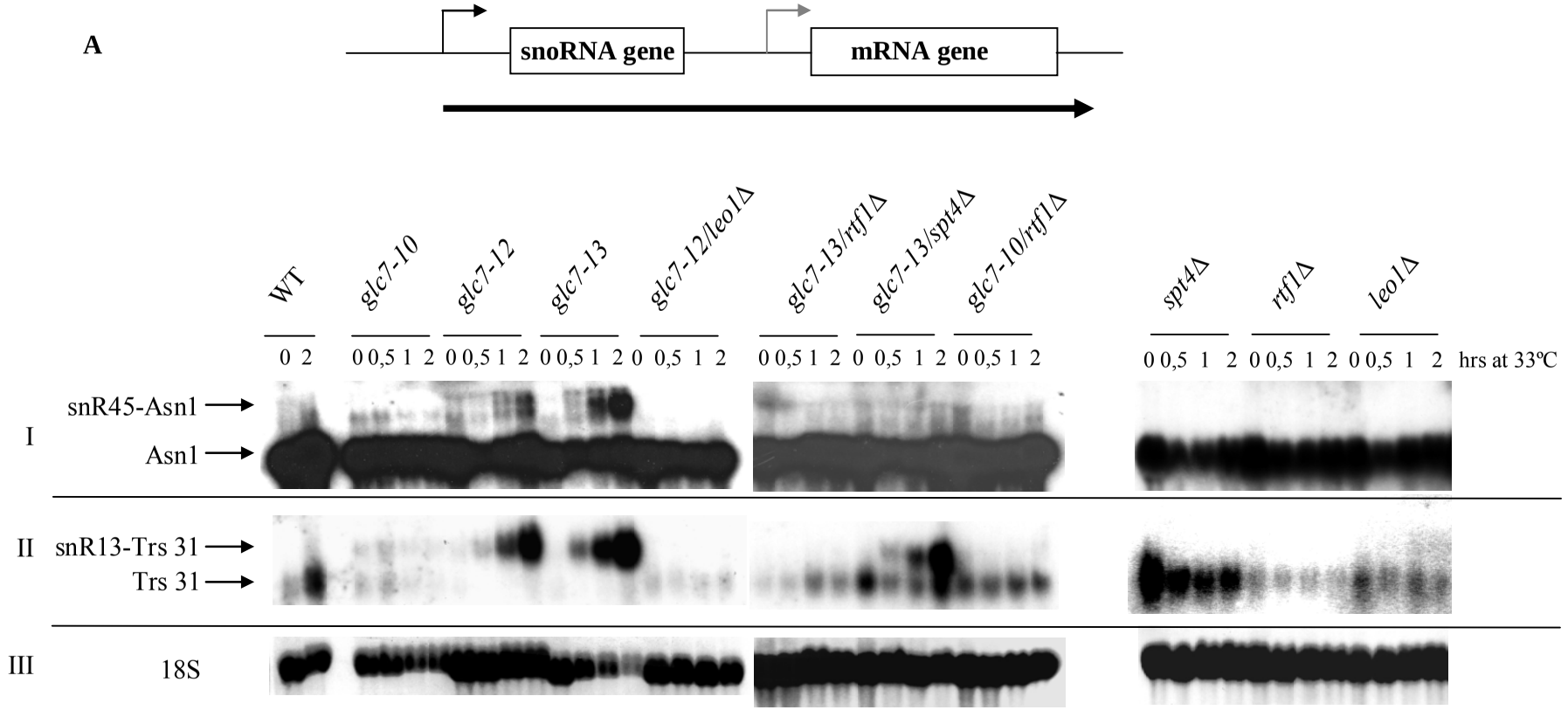
Strikingly, the read-through observed in the *glc7* mutants was completely suppressed in the *glc7-12/leoΔ* and *glc7-13/rtfΔ* double mutants and in part in the double mutant *glc7-13/spt4Δ,* indicating that the snoRNA transcription termination defect of the single glc7 mutants is suppressed by deletion of the elongation factors Leo1p, Rtf1p and to some extent Spt4p in the double mutants.

Accordingly, similar suppression effects were observed when we probed the double mutants with Nrd1 or Mrpl7, indicating that also in this case transcription termination was restored (Fig. 4.6 b). In addition, accumulation of the Nrd1 message itself, a phenotype characteristic of mutated factors of the Nrd1-dependent pathway of snoRNA transcription termination, is not observed in the double mutants *glc7-12/leo1Δ*, *glc7-13/rtf1Δ* and *glc7-10/rtf1Δ,* suggesting that autoregulation of Nrd1

expression is restored as well. However, although the double mutant *glc7-13/spt4Δ* has significantly less read-through product compared to the *glc7-13* mutant, it shows a strong accumulation of the Nrd1 message. This suggests that the read-through phenotype and the Nrd1 accumulation phenotype seem to be uncoupled in *glc7-13/spt4Δ*.

Altogether, we conclude that suppression of the snoRNA transcription termination defect in the double mutants could be one of the possible reasons why growth of *glc7-12* and *glc7-13* is restored in the double mutants *glc7-12/leoΔ* and *glc7-13/rtfΔ* at 33°C.

Moreover, we propose that the Paf1 complex subunits Rtf1 and Leo1 and to a much lower degree Spt4 can act as anti-terminators in snoRNA transcription. This supplements recent results for Spt4 as an anti-terminator in the transcription of protein coding genes (Kaplan et al., 2005). However, our results are in conflict with data that showed that members of the Paf1 complex have a transcription termination defect of snoRNAs (Sheldon et al., 2005; see discussion).


A
snoRNA gene
mRNA gene
WT
glc7-10
glc7-12
glc7-13
glc7-12/leo1Δ
glc7-13/rtf1Δ
glc7-13/spt4Δ
glc7-10/rtf1Δ
spt4Δ
rtf1Δ
leo1Δ
0 2
0 0,5 1 2
hrs at 33ºC
I
snR45-Asn1
Asn1
II
snR13-Trs 31
Trs 31
III
18S

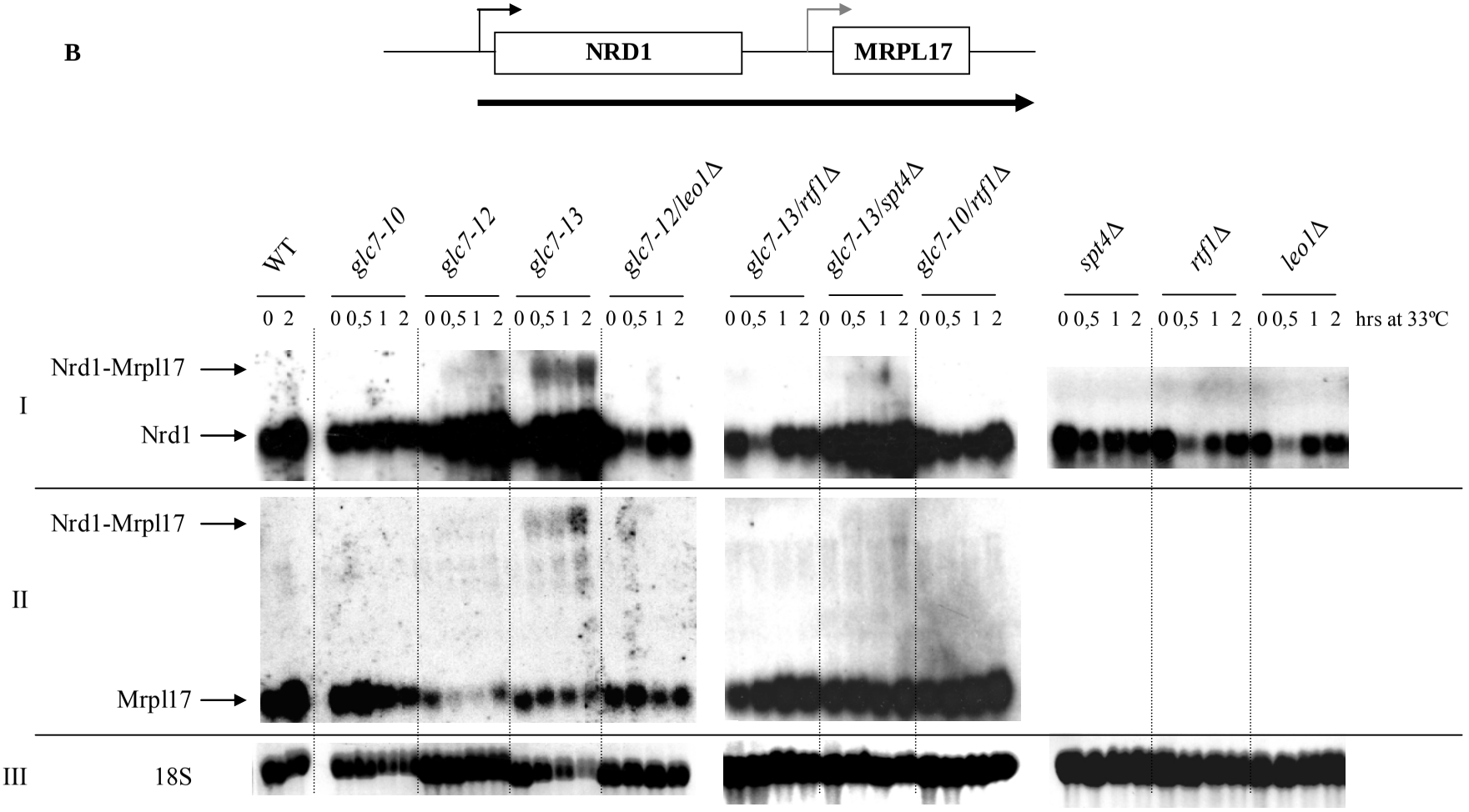

B
NRD1
MRPL17
WT
glc7-10
glc7-12
glc7-13
glc7-12/leo1Δ
glc7-13/rtf1Δ
glc7-13/spt4Δ
glc7-10/rtf1Δ
spt4Δ
rtf1Δ
leo1Δ
0 2
0 0,5 1 2
hrs at 33ºC
Nrd1-Mrpl17
Nrd1
I
Nrd1-Mrpl17
II
Mrpl17
III
18S

Figure 4.6. Accumulation of snoRNA-mRNA and Nrd1-Mrpl17 tandem transcripts is suppressed in double mutants of Glc7 with transcription elongation factors.
A) The drawing on top illustrates the genomic arrangement of the analysed genes and the relative direction of transcription. Northern analysis (panel I – III) of 20µg total RNA extracted from WT, *glc7-10*, *glc7-12*, *glc7-13,* glc7-13 x rtf1Δ, glc7-12 x leo1Δ, glc7-10 x rtf1Δ, glc7-13 x spt4Δ, spt4Δ, rtf1Δ and leo1Δ cells following growth in YPD medium at 25º C or after shift to 33º C for the indicated times. Blots were probed against Asn1 (panel I), Trs31 (panel II) and 18S (panel III; loading control). Migration of RNAs is indicated on the left. B) The drawing on top illustrates the genomic arrangement of the NRD1-MRPL17 locus and the relative direction of transcription. Northern analysis (panel I – III) of 20µg total RNA extracted from strains indicated on top following growth in YPD medium at 25º C or after shift to 33º C for the indicated times. Blots were probed against Nrd1 (panel I), Mrpl17 (panel II) and 18S (panel III; loading control). Migration of RNAs is indicated on the left.

4.4 Discussion

Transcription elongation is a highly regulated process. Several transcription elongation factors have been identified that either stimulate or inhibit this process. Here, we propose that also Glc7p, an essential protein phosphatase, is involved in the regulation of transcription elongation. First indications came from TRO experiments, which revealed that poly(A)-dependent pausing was enhanced in *glc7-12* mutants. Further investigations indicated that the *glc7-13* and *glc7-5* mutant strains were sensitive to 6AU, which is diagnostic of defects in transcription elongation. Interestingly, we also observed a strong resistance to 6AU for the *glc7-10* mutant strain. Furthermore, Glc7p genetically interacted with the elongation factors Spt4, Rtf1p and Leo1. Double mutants of Glc7p and these factors partially restored growth of Glc7p at semi-restrictive and restrictive temperatures. In addition, the snoRNA transcription termination defect, observed in *glc7* mutant strains was suppressed in these double mutants. This suggests that Glc7p is in fact a transcription termination factor that could be involved in the coupling of transcription elongation and termination.

What function could Glc7p have in transcription elongation?

Poly(A)-dependent pausing is a phenomenon thought to facilitate transcription termination of pre-mRNAs (Orozco et al., 2002). From our observation that Glc7p promotes poly(A)-dependent pausing we would expect that Glc7p is also required for transcription termination of pre-mRNAs. However, we did not observe a defect in transcription termination of pre-mRNAs in *glc7* mutant strains. Perhaps poly(A)-dependent pausing can also be considered as a checkpoint that ensures that 3' end processing takes place efficiently before RNAP II terminates transcription. This idea is supported by the observation that transcriptional pausing activates polyadenylation in a coupled *in vitro* transcription-polyadenylation system (Yonaha and Proudfoot, 1999). Similarly, promoter-proximal pausing was shown to be an important checkpoint to confirm that capping at the 5' end of the pre-mRNA has occurred before RNAP II is released from this pause site. The phosphatase activity of Glc7p is required for proper polyadenylation of pre-mRNAs *in vitro* and *in vivo* (see Chapter 2; (He and Moore, 2005). It is possible that the function of Glc7p to promote pausing

of RNAP II close to the poly(A) site is connected to its role in polyadenylation. Pausing could ensure that CPF was properly released from RNAP II and was able to attach to the poly(A) site. If pausing was blocked, CPF would only bind inefficiently to the poly(A) site resulting in a lack of polyadenylated mRNAs.

One other possibility is that cleavage is reversible and commitment for 3' end processing is only achieved after poly (A) tail formation has started. The poly(A)-dependent pausing checkpoint would then only be released from RNAP II after polyadenylation had started.

Chromatin represents a major block for RNAP II transcription and therefore has to be remodeled during transcription elongation (Sims et al., 2004a). Histone methylation at specific lysines appears to play a role in establishing both short-and long-term transcriptional regulation (Sims et al., 2004a). Transcriptionally active domains in chromatin are typically associated with methylation at H3-K4, H3-K36 and H3-K79 (Sims et al., 2003). These lysines are methylated by specific methyl transferases called Set1p and Set2p. The PAF complex has been shown to be required for the recruitment of Set1p and Set2p and also for correct H3-K4, H3-K36 and H3-K79 methylation (Krogan et al., 2003a; Krogan et al., 2003b; Ng et al., 2003b; Wood et al., 2003). In addition, the PAF subunit Rtf1p is specifically required for the recruitment of the H2B monoubiquitinylation enzyme and therefore for monoubiquitinylation of H2B (Wood et al., 2003). This reaction was shown to be a prerequisite for H3-K4 and H3-K79 methylation (Ng et al., 2003a; Wood et al., 2003). In conclusion, all actions of the PAF complex create a transcriptionally permissive chromatin structure. Glc7p is a factor that promotes transcription termination of snoRNAs (Chapter 3 and 4). The genetic interactions of Glc7p with Rtf1p that result in a suppression of the snoRNA transcription termination defect of Glc7p may represent antagonistic interactions: Rtf1 promotes an "open" chromatin structure that facilitates transcription elongation but in contrast does not support transcription termination. Glc7p on the other hand could counteract Rtf1p by regulating the activity of Rtf1p such that chromatin assumes a transcriptionally repressed state. The repressed chromatin state could act as a barrier for RNAP II forcing it to terminate. Therefore, Glc7p might be involved in coupling transcription elongation to termination.

The interactions we observed between Glc7p and Spt4p were subtle. This might indicate that Glc7p's function in transcription elongation does not completely overlap with the one of Spt4p in the process. Thus, Glc7p could rather be a transcription

termination factor that influences the activity of elongation factors such as Spt4p in the termination region. Because the activity of Spt4p was shown to be important for transcription of long GC rich regions of protein coding genes, Spt4p might not be so essential for transcription of the rather short snoRNA genes. Therefore, the interaction of Glc7p with Spt4p could be very limited (see below). It was reported that *spt4Δ* mutant strains show reduced levels of read-through at the poly(A) site of Gal 10, which may reflect an important role for this factor in promoting transcriptional read-through of cryptic or weak polyadenylation signals within coding regions (Kaplan et al., 2005). We observed that read-through of snoRNA terminators was reduced in a glc7/spt4Δ double mutant, indicating that Spt4p is required for promoting read-through of snoRNA terminators as well. Recently, it has been reported that at least some snoRNA terminators are bipartite in sequence, one motif having similarity to sequences required for 3' end processing of pre-mRNAs including the poly(A) site (Steinmetz et al., 2006). An interesting implication of this is that Spt4p could also promote read-through of these cryptic poly(A) sites in snoRNA terminators to prevent polyadenylation of snoRNAs.

5 Microarray analysis of *glc7-12*

5.1 Summary

The catalytic subunit of *Saccharomyces cerevisiae* type 1 protein phosphatase is encoded by the essential gene GLC7. Glc7p is regulating diverse cellular processes. We sought to identify new functions of Glc7p by a whole genome analysis using Affymetrix microarrays. To this end, we compared the expression profile of the temperature-sensitive *glc7-12* yeast mutant and the WT strain at the non-permissive temperature of *glc7-12*. By this approach, we found that Glc7p regulates transcription of ribosomal protein and also Ribi genes, which encode factors required for ribosome biogenesis. In addition, comparison of the expression profiles gave us some first indications that Glc7p could be involved in snoRNA transcription termination.

5.2 Introduction

Protein phosphatase type 1 (PP1) is a serine/threonine-specific phosphatase that is encoded by the essential gene GLC7 in *S. cerevisiae* (Feng et al., 1991). Glc7p is involved in diverse cellular processes ranging from glycogen accumulation and carbohydrate metabolism to translation initiation, transcription, mitosis, meiosis and histone modifications (Stark, 1996). The GLC7 gene encodes the catalytic subunit of PP1. The catalytic activity of Glc7 is controlled by association with a range of regulatory or targeting subunits that influence its intracellular localization and substrate specificity. There is evidence that almost all Glc7p is present in complexes with such proteins (Bollen, 2001). Understanding the many cellular roles of Glc7p therefore requires the identification and characterization of all Glc7p regulators. In the past, a number of proteins that interact with and regulate Glc7p in specific processes have been identified by genetical and biochemical studies (Bollen, 2001; Stark, 1996). These proteins include Sds22p, Reg1p, Gac1p, Glc8 and Ypi1p (Garcia-Gimeno et al., 2003; MacKelvie et al., 1995; Nigavekar et al., 2002; Tu and Carlson, 1995; Wu et al., 2001). Genome-wide approaches in yeast using the two-hybrid system and systematic analysis of protein complexes have implicated many other potential Glc7p interacting proteins (Gavin et al., 2006; Gavin et al., 2002; Ito et al., 2001; Ito et al., 2000; Uetz et al., 2000). In most cases however, the interactions with Glc7p have not been confirmed by independent approaches. The two-hybrid studies and analyses of

protein complexes were concentrated on finding direct protein-protein interactions that could help to find new functions for Glc7p. We sought to explore novel cellular functions of Glc7p by a different approach. We applied microarray analysis to study the expression pattern of genes in cells that lack Glc7 phosphatase activity. We thereby compared which genes were differentially expressed in *glc7-12*, a temperature-sensitive Glc7 allele, at the restrictive temperature compared to the WT. The rationale behind these experiments was that specific changes in the expression pattern could allow to draw conclusion as to which cellular process might have caused these changes. Many signaling pathways communicate extracellular and intracellular conditions to the cell nucleus, ultimately leading to a transcriptional response. Signaling pathways are often regulated by phosphorylation and dephosphorylation of pathway components. Also transcription factors can be regulated by phosphorylation. Interfering with this regulation by loss of a phosphatase activity should alter the expression of target genes of these signaling cascades. The target genes can therefore reveal the signaling pathway which was interupted by loss of the phosphatase activity. Thus, alterations in the expression pattern could pinpoint at previously non-identified novel cellular functions and targets of Glc7p. Microarray analysis also has been applied by several laboratories to identify defects in snoRNA transcription termination (see Chapter 3; Ganem et al., 2003; Steinmetz et al., 2001).

Our previous work with the *glc7-12* allele indicated that Glc7p is required for snoRNA transcription termination (Chapter 3) and transcription elongation (Chapter 4). Here, we show that microarray analysis of *glc7-12* identified a new role for Glc7p in the regulation of ribosomal protein and Ribi gene transcription (Chapter 5 and 6). In addition, this microarray analysis was helpful to detect a snoRNA transcription termination defect in *glc7* mutants (Chapter 3).

5.3 Results

5.3.1 Experimental Design

The aim of this large-scale microarray study was to identify genes that show differential expression in the mutant strain *glc7-12* at 37º C compared to the WT.

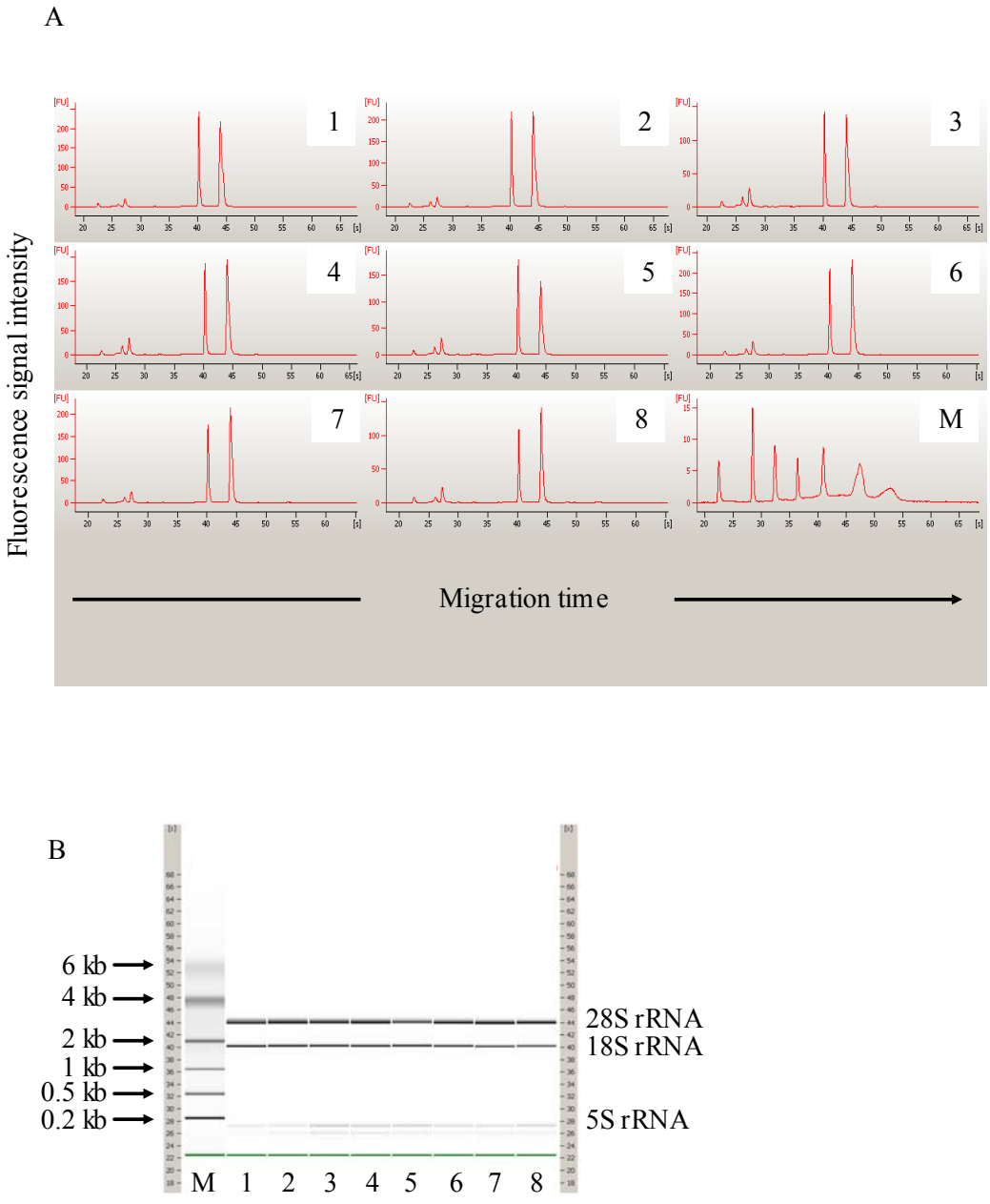
The restrictive temperature for the ts mutant strain *glc7-12* is 37º C, so Glc7p will be non-functional at this temperature. This leads to changes in numerous cellular processes that are regulated by Glc7p and will ultimately also lead to changes in gene expression since stress is known to elicit a transcriptional stress response. Therefore, new functions for Glc7p could be revealed by looking at the specific stress response and that might give a clue to which cellular process it connects. Glc7p could in addition directly regulate transcription factors. Their misregulation then would be apparent in an altered gene expression.

For this purpose High Density Oligonucleotide Microarrays (S98 GeneChips) from Affymetrix were used. The arrays contain approximately 6400 ORFs and the transcript levels of these can be measured simultaneously in one experiment. To this end, total RNA is extracted from cells which is converted to cDNA and amplified by PCR. Finally, cRNA is transcribed from the amplified cDNA. The cRNA is subsequently fragmented to ensure better hybridization to the probe sets on the microarray chip and labeled with a fluorophor to allow later detection. A probe set refers to 16 oligonucleotides present on the microarray that are complementary to the 3' UTRs of one given transcript.

Because RNA degradation could cause data to be less reproducible and make it more difficult to detect low-abundance transcripts the RNA quality is tested before the use in the microarray experiment.

For this, extracted total RNA is purified on RNeasy Columns and analyzed with the Agilent Bioanalyzer RNA Chip (see materials and methods). Figure 5.1 a shows a compilation of the RNA profiles obtained from the Bioanalyzer and demonstrates that our total RNA was of very high quality, indicated by the two sharp peaks which represent 28S and 18S rRNA. A broad peak or additional peaks would denote RNA degradation. The data of the Bioanalyzer were also processed and displayed as a virtual gel (Fig. 5.1 b). Again, sharp bands indicated that there was no degradation of the RNA.

In addition, cRNA target molecule preparations had to be tested with the Agilent Bioanalyzer RNA Chip (Fig. 5.1 c). Ideally, the cRNA preparation should reflect the average size of the messenger RNAs with around 1500 to 2000 nucleotides on average. The computed virtual gel shows that target cRNA was successfully generated with a high concentration and on expected average size (Fig. 5.1 d). In conclusion, total RNA and cRNA preparations made from all strains and conditions were of premium quality.



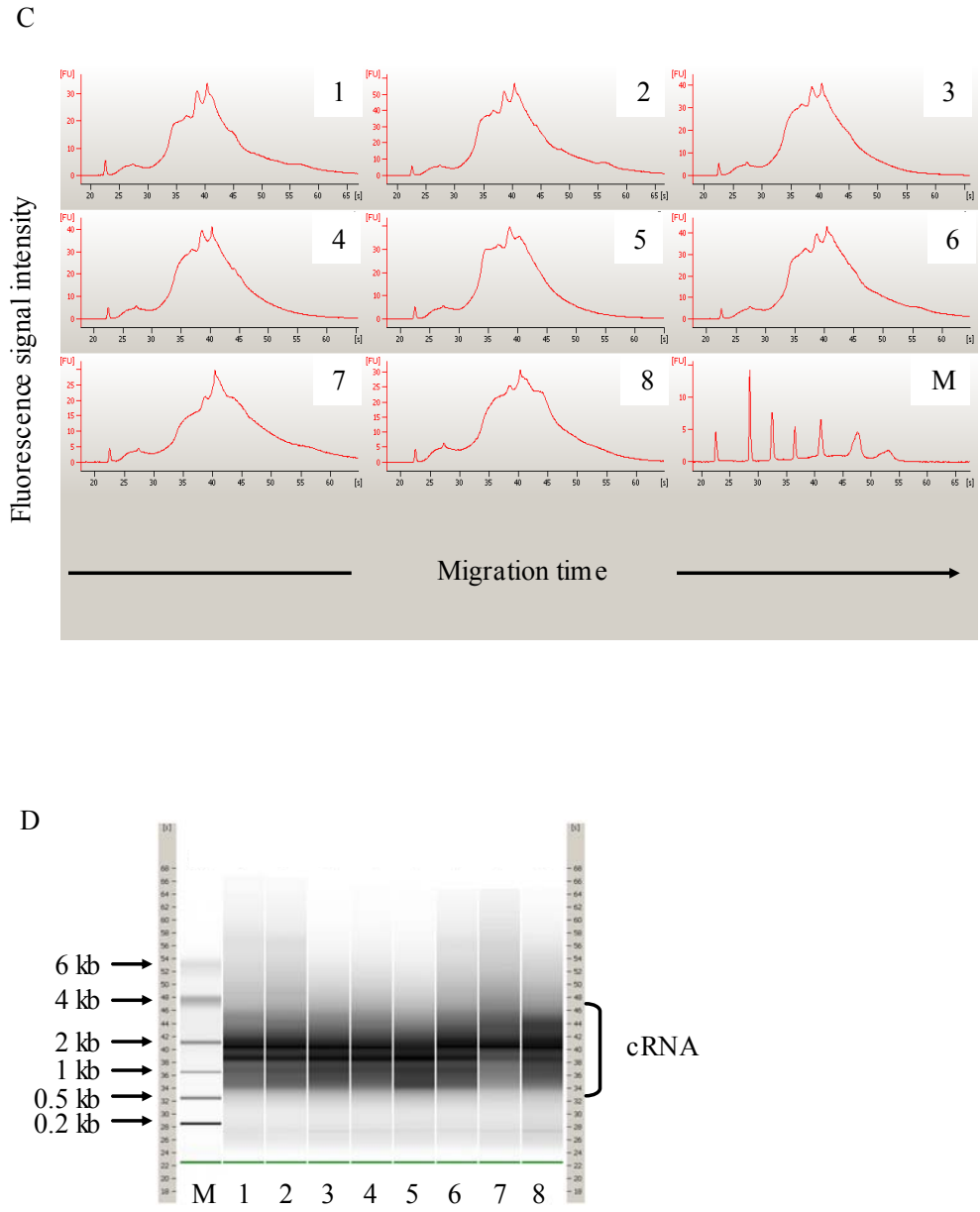


Figure 5.1. Total RNA and cRNA target quality assessment
Purified total RNA from (1) WT at RT; (2) WT at RT, replicate; (3) WT at 37º C; (4) WT at 37º C, replicate; (5) *glc7-12* at RT; *glc7-12* at RT, replicate; *glc7-12* at 37º C; and *glc7-12* at 37º C, replicate cells were analyzed for their concentration and overall length by using the Agilent Bioanalyzer RNA Chip. Fluorescence intensities measured with total RNA and cRNA target molecules are plotted against migration time in seconds as indicated in A) for total RNA and C) for cRNA. The peaks corresponding to the molecular weight RNA ladder are displayed (M). Virtual gels of the total RNA and cRNA samples as indicated are shown in B) and D), respectively. M, molecular weight markers.

5.3.2 Identification of differentially expressed transcripts in glc7-12 versus WT cells

To identify genes that were differentially expressed in the mutant strain *glc7-12* versus the WT we had to investigate four different conditions, namely both WT and *glc7-12* at RT and after shift to 37º C for 45 minutes (WT at RT, WT at 37º C, *glc7-12* at RT, *glc7-12* at 37º C).

Each condition is represented by one microarray experiment. To ensure reproducibility and therefore a high significance of the results, we performed each microarray experiment in duplicate. By comparing the outcomes of these eight experiments we could extract those genes that differed in transcript levels between RT and 37ºC and that clearly differed at 37ºC between the WT and *glc7-12* mutant to exclude mere temperature effects. To filter for these genes we applied a p-value of < 0.001. This means that the probability that the results observed could have occurred by chance is less than 1‰. With these criteria we found 826 genes to be differentially expressed of which 417 were upregulated and 409 were downregulated (Christa Wiederkehr, personal communication).

Before gene filtering, the data had to be preprocessed to enable comparison of data obtained from different microarray chips. Preprocessing includes background correction, normalization, calculation of the expression measure and a quality control.

Normalization is a process to reduce the variation of non-biological origin across arrays. Non-biological variations include variation introduced during sample preparation, manufacturing and processing of the arrays (labeling, hybridization and scanning). One way to demonstrate differences in the distribution of intensities between distinct microarray experiments can be done by a Box plot. A Box plot visualizes the need to normalize the data of different experiments. When we analyzed our unnormalized data in a Box plot, we observed very little alteration between the different microarrays, indicating that non-biological variations were reduced to a minimum (Fig. 5.2 a; Christa Wiederkehr, personal communication).

Another test to analyze whether expression levels differ within replicates is the Scatter blot. For both the WT and *glc7-12* at RT and 37º C the scatter plot matrix indicated that the duplicates were highly reproducible (Fig. 5.2 b; Christa Wiederkehr, personal communication).

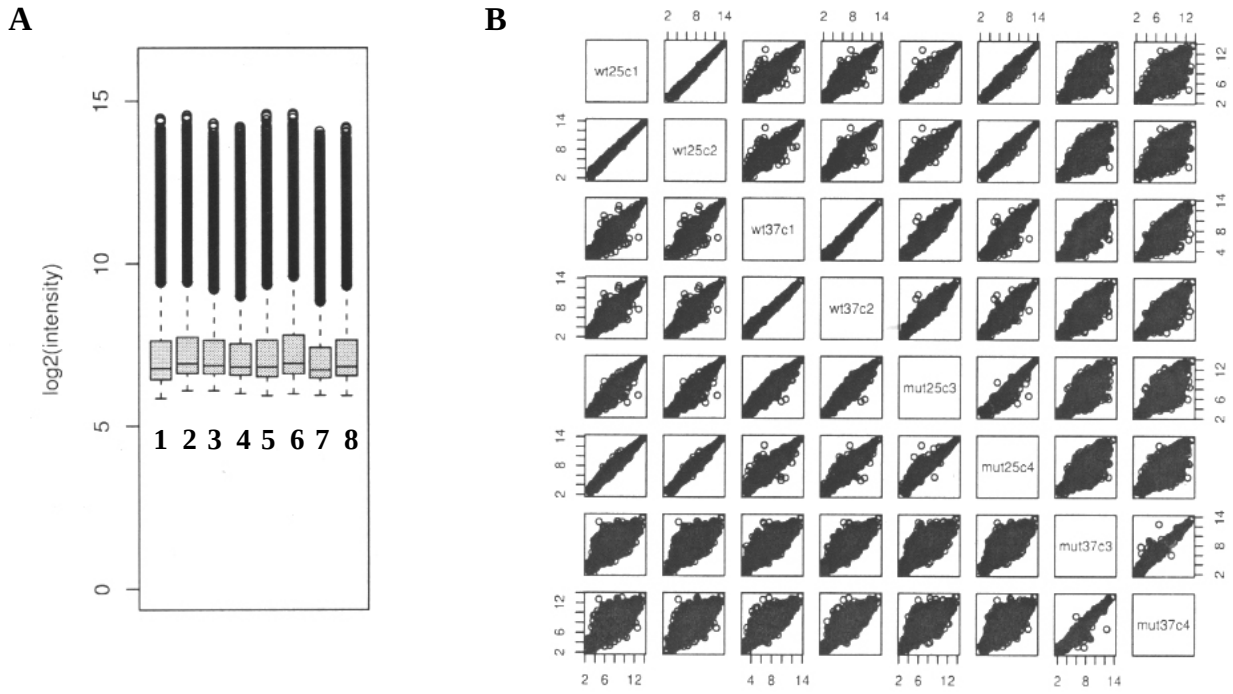


Figure 5.2. Comparing gene expression levels of probes sets within replicates and different strains.
Box plots in A) display the overall distribution of observed changes in signal intensities obtained from (1) WT at RT; (2) WT at RT, replicate; (3) WT at 37º C; (4) WT at 37º C, replicate; (5) *glc7-12* at RT; *glc7-12* at RT, replicate; *glc7-12* at 37º C; and *glc7-12* at 37º C, replicate cells. Shown are the absolute log2 fold change values. The features of clear outliers were manually inspected on the corresponding DAT image file to exclude artifacts. The scatter plot matrix in B) summarizes a systematic comparison of the log2-scaled expression signals from transcripts among all samples as indicated.

5.3.3 Clustering gene expression profiles

In order to obtain a better understanding of functional classes of the genes we identified, the two groups of upregulated or downregulated genes were further subdivided into three groups each, such that the 826 genes were finally grouped into 6 clusters (Fig. 5.3; Christa Wiederkehr, personal communication). This way, genes in one cluster are more closely related to each other than to genes of a different cluster. Consequently, clustering can divide genes into functional classes on the basis of their expression pattern.

In addition, we used the FatiGo Gene Ontology tool to analyze our data. With the help of this program the genes of each Cluster are assigned to groups according to their biological processes, molecular functions or as part of a cellular component.

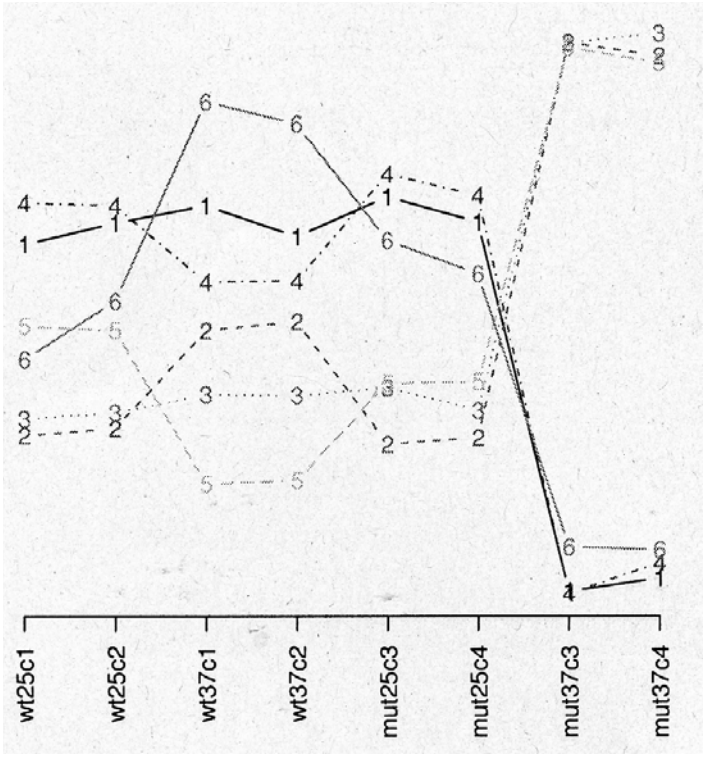


Figure 5.3. Typical expression pattern of genes in the six clusters
Numbers in graph indicate the respective cluster. Strains and conditions used as indicated below the graph.

5.3.4 Microarray results

5.3.4.1 Cluster 1

Cluster 1 contains 159 ORFs with downregulated expression in *glc7-12* mutant strains (see Table 5.1). From the 159 ORFs 22 are uncharacterized ORFs and 3 are not annotated yet.

Interestingly, this cluster contains 25 genes of the Ribi regulon (PRS3, URA7, URA1, AAH1, URK1, DBP8, YBR267W, NOG1, PRP43, ARX1, NOC4, PNO1, KRR1, UTP15, NOC3, NSA1, DIM1, DBP6, RPB8, RPC34, RPA49, MES1, WRS1, TRM82, GCD10) equal to 44% of all Ribi genes detected in our microarray analysis. The Ribi regulon consists of 236 genes that have functions in ribosome biogenesis (Jorgensen et al., 2004). Other Ribi genes were found in Cluster 4 and 6 (for more information see also Chapter 6).

We also found three ribosomal protein (RP) genes (RPL8A, RPS26A, RPL21B) in cluster 1. In addition a large amount of RP genes are found in cluster 4. It is likely that

cluster 1 and 4 are regulated in part by the same set of transcription factors (see Cluster 4).

Table 5.1

no.	systematic name	gene name	description
1	YLR295C	ATP14	ATP synthase subunit h
2	YLR292C	SEC72	protein involved in membrane protein insertion into the ER
3	YLR172C	DPH5	diphthamide biosynthesis
4	YLR179C		Hypothetical ORF
5	YLR154C	RNH203	Ribonuclease H2 subunit
6	YLR056W	ERG3	C-5 sterol desaturase
7	YLR063W		Hypothetical ORF
8	YLR073C		Hypothetical ORF
9	YLR002C	NOC3	Protein that forms a nuclear complex with Noc2p that binds to 66S ribosomal precursors to mediate their intranuclear transport; also binds to chromatin to promote the association of DNA replication factors and replication initiation
10	YLR008C	PAM18	Constituent of the mitochondrial import motor associated with the presequence translocase, along with Ssc1p, Tim44p, Mge1p, and Pam16p; stimulates the ATPase activity of Ssc1p to drive mitochondrial import; contains a J domain; DnaJ-like protein, cochaperone
11	YKR039W	GAP1	general amino acid permease
12	YKL029C	MAE1	malic enzyme
13	YKL128C	PMU1	phosphomutase homolog
14	YKL165C	MCD4	Required for GPI anchor synthesis
15	YKL216W	URA1	dihydroorotate dehydrogenase
16	YKL214C	YRA2	Yeast RNA Annealing protein
17	YJR112W	NNF1	Spindle pole protein, required for accurate chromosome segregation
18	YJR057W	CDC8	thymidylate kinase
19	YJR024C		Hypothetical ORF
20	YJR003C		Hypothetical ORF
21	YJL069C		U3 snoRNP protein|U3 snoRNA associated protein
22	YJL061W	NUP82	82 kDa protein, with putative coiled-coil domain, has carboxy-terminal domain, containing heptad repeats, that binds Nsp1p|nuclear pore complex subunit|nucleoporin
23	YJL081C	ARP4	54.8 kDa protein|actin related protein
24	YJL131C		Hypothetical ORF
25	YJL186W	MNN5	golgi alpha-1,2-mannosyltransferase (putative)
26	YJL181W		Hypothetical ORF
27	YAL033W	POP5	RNase MRP subunit (putative)|RNase P integral subunit
28	YAL036C	FUN11	Protein of unknown function, has similarity to Xenopus GTP-binding protein DRG
29	YIR026C	YVH1	protein tyrosine phosphatase induced by nitrogen starvation
30	YIL020C	HIS6	phosphoribosyl-5-amino-1-phosphoribosyl-4-imidazolecarboxiamide isomerase
31	YIL016W	SNL1	18.3 kDa integral membrane protein
32	YIL044C	AGE2	ARF GAP with effector function(s)
33	YIL094C	LYS12	homo-isocitrate dehydrogenase
34	YHR197W	RIX1	RIbosome eXport
35	YHR179W	OYE2	NAPDH dehydrogenase (old yellow enzyme), isoform 2
36	YHR154W	RTT107	Regulator of Ty1 Transposition; Establishes Silent Chromatin

37	YHR163W	SOL3	weak multicopy suppressor of los1-1
38	YHR169W	DBP8	dead box protein
39	YHR149C	SKG6	Hypothetical ORF, found in the bud tip and bud neck, potential Cdc28p substrate; Skg6p interacts with Zds1p and Zds2p
40	YHR098C	SFB3	similar to SEC24
41	YHR085W	IPI1	Involved in Processing ITS2
42	YHR049W	FSH1	Serine hydrolase that localizes to both the nucleus and cytoplasm; sequence is similar to Fsh2p and Fsh3p
43	YHR031C	RRM3	DNA helicase
44	YHR007C	ERG11	cytochrome P450 lanosterol 14a-demethylase
45	YHR013C	ARD1	N alpha-acetyltransferase major subunit|complexes with Nat1p
46	YHL011C	PRS3	ribose-phosphate pyrophosphokinase
47	YHL033C	RPL8A	ribosomal protein L8A (rp6) (YL5) (L4A)
48	YGR264C	MES1	methionine-tRNA ligase
49	YGR234W	YHB1	flavohemoglobin
50	YGR147C	NAT2	N alpha-acetyltransferase
51	YGR128C	UTP8	U3 snoRNP protein
52	YGR090W	UTP22	U3 protein
53	YGR060W	ERG25	C-4 sterol methyl oxidase
54	YGR049W	SCM4	Protein that suppresses ts allele of CDC4 when overexpressed
55	YGR017W		Hypothetical ORF
56	YGL027C	CWH41	glucosidase I
57	YGL120C	PRP43	RNA helicase
58	YGL111W	NSA1	ribosome biogenesis
59	YGL189C	RPS26A	ribosomal protein S26A
0	YGL200C	EMP24	type I transmembrane protein
61	YGL246C	RAI1	Product of gene unknown
62	YFL017C	GNA1	glucosamine-phosphate N-acetyltransferase
63	YFL026W	STE2	alpha-factor pheromone receptor|seven-transmembrane domain protein
64	NC-001137	NA	NA
65	YER166W	DNF1	Potential aminophospholipid translocase
66	YER072W	VTC1	S. pombe Nrf1p homolog (97% identical in predicted amino acid sequence)
67	YER019C-A	SBH2	Sbh1p homolog
68	YDR465C	RMT2	arginine methyltransferase
69	YDR400W	URH1	uridine nucleosidase (uridine ribohydrolase); EC 3.2.2.3
70	YDR395W	SXM1	karyopherin beta family member
71	YDR368W	YPR1	homologous to the aldo-keto reductase protein family
72	YDR345C	HXT3	low affinity glucose transporter
73	YDR352W		Hypothetical ORF
74	YDR292C	SRP101	signal recognition particle receptor - alpha subunit
75	YDR257C	RMS1	Transcription regulator
76	YDR226W	ADK1	adenylate kinase
77	YDR165W	TRM82	Transfer RNA methyltransferase
78	YDR101C	ARX1	Shuttling pre-60S factor; involved in the biogenesis of ribosomal large subunit biogenesis; interacts directly with Alb1; responsible for Tif6 recycling defects in absence of Rei1; associated with the ribosomal export complex
79	YDR056C		Hypothetical ORF
80	YDL063C		Hypothetical ORF
81	YDL095W	PMT1	dolichyl phosphate-D-mannose:protein O-D-mannosyltransferase
82	YDL167C	NRP1	Protein of unknown function, rich in asparagine residues

83	YDL236W	PHO13	p-nitrophenyl phosphatase
84	YCR087C-A	LUG1	Hypothetical ORF
85	YCL025C	AGP1	amino acid permease
86	YCL059C	KRR1	Involved in cell division and spore germination
87	YBR267W	REI1	Cytoplasmic pre-60S factor; required for the correct recycling of shuttling factors Alb1, Arx1 and Tif6 at the end of the ribosomal large subunit biogenesis; involved in bud growth in the mitotic signaling network
88	YBR261C		Putative S-adenosylmethionine-dependent methyltransferase of the seven beta-strand family
89	YBR229C	ROT2	glucosidase II
90	YBR205W	KTR3	alpha-1,2-mannosyltransferase (putative)
91	YBR205W	KTR3	alpha-1,2-mannosyltransferase (putative)
92	YBR206W	NA	NA
93	YBR106W	PHO88	May be a membrane protein involved in inorganic phosphate transport and regulation of Pho81p function
94	YBR085W	AAC3	ADP/ATP translocator
95	YBR092C	PHO3	acid phosphatase
96	YBR017C	KAP104	karyopherin beta 2
97	YBL039C	URA7	CTP synthase
98	YBL035C	POL12	DNA polymerase alpha-primase complex B subunit
99	YBL057C	PTH2	peptidyl-tRNA hydrolase
100	YBL081W		Hypothetical ORF
101	YPR144C	NOC4	U3 snoRNP protein
102	YPR138C	MEP3	NH4+ transporter
103	YPR060C	ARO7	chorismate mutase
104	YPR062W	FCY1	cytosine deaminase
105	YPR036W	VMA13	vacuolar H(+) ATPase V1 sector 54 kDa subunit
106	YPL079W	RPL21B	ribosomal protein L21B
107	YPL094C	SEC62	ER protein translocation apparatus membrane component
108	YPL093W	NOG1	homologs identified in human and Trypanosoma brucei|nucleolar G-protein (putative)
109	YPL103C	FMP30	Hypothetical ORF
110	YPL144W		Hypothetical ORF
111	YPL207W	TWY1	Hypothetical ORF
112	YPL244C	HUT1	similar to UDP-galactose transporter
113	YPL266W	DIM1	dimethyladenosine transferase
114	YOR375C	GDH1	NADP-specific glutamate dehydrogenase
115	YOR307C	SLY41	chloroplast phosphate transporter homolog
116	YOR251C		catalyzes transfer of the sulfane atom of thiosulfate to cyanide to form sulfite and thiocyanate
117	YOR212W	STE4	G protein beta subunit|coupled to mating factor receptor
118	YOR224C	RPB8	16 kDa RNA polymerase subunit (common to polymerases I, II and III)
119	YOR145C	PNO1	Associated with Nob1
120	YOR101W	RAS1	ras homolog
121	YOR112W		Hypothetical ORF
122	YOR118W		Hypothetical ORF
123	YOR039W	CKB2	protein kinase CK2, beta' subunit
124	YOL007C		Appears to be a structural component of the chitin synthase 3 complex
125	YOL022C	CSI2	Protein required for cell viability
126	YOL080C	REX4	RNA EXonuclease; member of 3'->5' exonuclease family. See Moser et al. 1997 Nucleic acids Res. 25:5110-5118
127	YOL097C	WRS1	tryptophan-tRNA ligase
128	YOL143C	RIB4	6,7-dimethyl-8-ribityllumazine synthase (DMRL synthase)
129	YOL136C	PFK27	6-phosphofructo-2-kinase
130	YNR046W	TRM112	Subunit of an adoMet-dependent tRNA methyltransferase

			(MTase) complex (Trm11p-Trm112p), required for the methylation of the guanosine nucleotide at position 10 (m2G10) in tRNAs; putative zinc binding subunit of other MTase-related proteins
131	YNR038W	DBP6	RNA helicase (putative)
132	YNR003C	RPC34	RNA polymerase III (C) 34 kDa subunit
133	YNR012W	URK1	uridine kinase
134	YNL026W	SAM50	Essential component of the Sorting and Assembly Machinery (SAM or Tob complex) of the mitochondrial outer membrane, which binds precursors of beta-barrel proteins and facilitates their outer membrane insertion; homologous to bacterial Omp85
135	YNL010W		Hypothetical ORF
136	YNL049C	SFB2	zinc finger protein (putative)
137	YNL062C	GCD10	RNA-binding protein|subunit of tRNA(1-methyladenosine) methyltransferase, along with Gcd14p
138	YNL111C	CYB5	cytochrome b5
139	YNL102W	POL1	DNA polymerase I alpha subunit p180
140	YNL119W	NCS2	Hypothetical ORF
141	YNL145W	MFA2	a-factor mating pheromone precursor
142	YNL141W	AAH1	adenine aminohydrolase (adenine deaminase)
143	YNL182C	IPI3	Protein required for cell viability; computational analysis of large-scale protein-protein interaction data suggests a possible role in assembly of the ribosomal large subunit; Involved in Processing ITS2
144	YNL248C	RPA49	RNA polymerase A 49 kDa alpha subunit
145	YNL246W	VPS75	Protein of unknown function involved in vacuolar protein sorting; detected in the nucleus
146	YNL289W	PCL1	G1 cyclin|associates with PHO85
147	YNL316C	PHA2	prephenate dehydratase
148	SNR76	NA	NA
149	YNL333W	SNZ2	Snooze: stationary phase-induced gene family
150	YNL323W	LEM3	membrane glycoprotein
151	YMR316C-B	NA	NA
152	YMR127C	SAS2	zinc finger protein
153	YMR093W	UTP15	U3 snoRNP protein
154	YML060W	OGG1	43 kDa 8-oxo-guanine DNA glycosylase
155	YML104C	MDM1	intermediate filament protein
156	YML125C	PGA3	Essential protein required for maturation of Gas1p and Pho8p, protein trafficking; GFP-fusion protein localizes to the endoplasmic reticulum; null mutants have a cell separation defect
157	YLR412W		Hypothetical ORF
158	YLR420W	URA4	dihydrooratase
159	YLR351C	NIT3	Nit protein, nitrilase superfamily member

5.3.4.2 Cluster 2

Cluster 2 contains 121 ORFs where expression is upregulated in *glc7-12* mutant strains (Table 5.2). Out of the 121 ORFs 27 are uncharacterized and 10 are not annotated yet.

Two interesting genes in this cluster are TPK2 and MKK1. TPK2 is one of three redundant catalytic subunits of cAMP-dependent protein kinase PKA (Toda et al., 1987b). We have seen mislocalization of TPK1, the other of the catalytic subunits, in *glc7-13* mutant cells (see Chapter 6). The cells might therefore try to counteract this process by producing more TPK2 to abide the function of PKA in the control of RP gene transcription.

MKK1 encodes the MAP kinase kinase (MEK). Signaling through the MAP kinase pathway also contributes to the regulation of RP gene transcription. Thus, both genes are involved in signaling cascades that can influence the transcription of RP genes.

Table 5.2

no.	systematic name	gene name	description
1	YLR312C		Hypothetical ORF
2	YLR054C	OSW2	Protein of unknown function proposed to be involved in the assembly of the spore wall
3	YLR031W		Hypothetical ORF
4	YLL026W	HSP104	heat shock protein 104
5	YKR098C	UBP11	ubiquitin-specific protease
6	YKL091C	SFH1	Putative homolog of Sec14p, which is a phosphatidylinositol/phosphatidylcholine transfer protein involved in lipid metabolism; localizes to the nucleus
7	YKL121W		Hypothetical ORF
8	YKL151C		Hypothetical ORF
9	YKL142W	MRP8	ribosomal protein
10	NC-001142	NA	NA
11	YJR096W		Protein with similarity to aldo-keto reductases
12	YJL017W	NA	NA
13	YJL082W	IML2	Similar to Ykr018p
14	YJL144W		Hypothetical ORF
15	YJL126W	NIT2	Nit protein, nitrilase superfamily member
16	YJL149W		Hypothetical ORF; has similarity to F-box proteins
17	YJL221C	FSP2	similar to maltase (alpha-D-glucosidase)
18	YAL028W	FRT2	Tail-anchored endoplasmic reticulum membrane protein, interacts with homolog Frt1p but is not a substrate of calcineurin (unlike Frt1p), promotes growth in conditions of high Na+, alkaline pH, or cell wall stress; potential Cdc28p substrate
19	YAL061W		putative polyol dehydrogenase
20	YBLWTAU1	NA,long terminal repeat	NA, Ty4 LTR
21	YLRWDELTA15	NA, long terminal repeat	NA ,Ty1 LTR
22	YIL045W	PIG2	30% identity to YER054C/GIP2
23	YIL117C	PRM5	pheromone-regulated membrane protein
24	YIL099W	SGA1	glucoamylase

25	YIL097W	FYV10	Function required for Yeast Viability on toxin exposure
26	YIL136W	OM45	45 kDa mitochondrial outer membrane protein
27	YIL153W	RRD1	Resistant to Rapamycin Deletion
28	YIL152W		Hypothetical ORF
29	YHR161C	YAP1801	Yeast Assembly Polypeptide, member of AP180 protein family, binds Pan1p and clathrin
30	YHR087W		Protein involved in RNA metabolism, one of two yeast homologs (with Sdo1p/Ylr022cp) of the human protein SBDS responsible for autosomal recessive Shwachman-Bodian-Diamond Syndrome, also conserved in Archaea
31	YHR096C	HXT5	hexose transporter
32	YHR009C		Hypothetical ORF
33	NC-001139	NA	NA
34	YGR142W	BTN2	Gene/protein whose expression is elevated in a btn1 minus/Btn1p lacking yeast strain.
35	YGR146C		Hypothetical ORF
36	YGR161C	RTS3	Hypothetical ORF, Putative component of the protein phosphatase type 2A complex
37	YGR127W		Hypothetical ORF
38	YGR138C	TPO2	Polyamine transport protein
39	YGR043C		Hypothetical ORF
40	YGL153W	PEX14	Peroxisomal peripheral membrane protein (peroxin) involved in import of peroxisomal matrix proteins
41	YGL158W	RCK1	Serine/threonine protein kinase
42	YGL156W	AMS1	alpha mannosidase
43	YGL248W	PDE1	3',5'-cyclic-nucleotide phosphodiesterase, low affinity
44	YFR017C		Hypothetical ORF
45	YFL014W	HSP12	heat shock protein 12
46	YER103W	SSA4	HSP70 family
47	YER101C	AST2	Protein involved in targeting of plasma membrane [H+]ATPase
48	YER079W		Hypothetical ORF
49	YER035W	EDC2	Enhancer of mRNA Decapping
50	YER015W	FAA2	acyl-CoA synthetase (fatty acid activator 2)
51	YDR479C	PEX29	peroxin
52	YDR425W	SNX41	sorting nexins Snx4p, Snx41p, and Snx42p mediate distinct retrieval pathways from endosomes.
53	YDR436W	PPZ2	serine-threonine phosphatase Z
54	YDR171W	HSP42	Similar to HSP26; expression is regulated by stress conditions
55	YDR124W		Hypothetical ORF
56	YDR069C	DOA4	ubiquitin isopeptidase
57	YDR074W	TPS2	trehalose-6-phosphate phosphatase
58	YDR043C	NRG1	binds to UAS-1 in the STA1 promoter and can interact with Ssn6p|transcriptional repressor
59	YDL085W	NDE2	Mitochondrial external NADH dehydrogenase, catalyzes the oxidation of cytosolic NADH; Nde1p and Nde2p are involved in providing the cytosolic NADH to the mitochondrial respiratory chain
60	YDL080C	THI3	alpha-ketoisocaproate decarboxylase
61	YDL134C	PPH21	serine-threonine protein phosphatase 2A
62	YDL169C	UGX2	Product of gene unknown
63	YDL233W		Hypothetical ORF
64	YDL223C	HBT1	Hub1 target
65	YCL044C	MGR1	Subunit, with Yme1p, of the mitochondrial inner membrane i-AAA protease complex, which is responsible for degradation of unfolded or misfolded mitochondrial gene products; required for growth of cells lacking the mitochondrial genome
66	NC-001134	NA	NA
67	NC-001134	NA	NA

68	YBR299W	MAL32	maltase
69	YBR214W	SDS24	Similar to S. pombe SDS23, suppresses DIS2, localized to the nucleus
70	YBR169C	SSE2	HSP70 family|SSE1 homolog
71	YBR182C	SMP1	Second MEF2-like Protein 1Transcription factor of the MADS (Mcm1p, Agamous, Deficiens, SRF) box family; closely related to RLM1
72	YBR148W	YSW1	Spore-specific protein
73	YBR149W	ARA1	D-arabinose dehydrogenase
74	YBR126C	TPS1	trehalose-6-phosphate synthase/phosphatase complex 56 kDa synthase subunit
75	YBR116C	NA	NA
76	YBL064C	PRX1	also called mTPx I, a mitochondrial isoform of thioredoxin peroxidase (EC 1.11.1.-)
77	YBL101C	ECM21	ExtraCellular Mutant
78	NC-001148	NA	NA
79	YPR154W	PIN3	[PSI+] induction
80	YPR079W	MRL1	Mannose 6-phosphate Receptor Like
81	YPR026W	ATH1	acid trehalase
82	YPR030W	CSR2	Nuclear protein with a potential regulatory role in utilization of galactose and nonfermentable carbon sources; overproduction suppresses the lethality at high temperature of a chs5 spa2 double null mutation
83	YPL166W	ATG29	Hypothetical ORF, Protein specifically required for autophagy; may function in autophagosome formation at the pre-autophagosomal structure in collaboration with other autophagy proteins
84	YPL156C	PRM4	pheromone-regulated membrane protein
85	YPL147W	PXA1	ABC transporter of long-chain fatty acids
86	YPL203W	TPK2	cAMP-dependent protein kinase catalytic subunit
87	YPL223C	GRE1	Induced by osmotic stress
88	YOR273C	TPO4	Polyamine transport protein
89	YOR275C	RIM20	Protein involved in proteolytic activation of Rim101p in response to alkaline pH; member of the PalA/AIP1/Alix family; interacts with the ESCRT-III subunits Snf7p, suggesting a relationship between the response to pH and multivesicular body formation
90	YOR231W	MKK1	MAP kinase kinase (MEK)
91	YOR227W		Hypothetical ORF
92	YOR173W	DCS2	Similar to mRNA decapping enzyme, HIT superfamily.
93	YOR185C	GSP2	GTP-binding protein|Gsp1p homolog
94	YOR162C	YRR1	transcription factor
95	YOR106W	VAM3	syntaxin family
96	YOR057W	SGT1	G2 allele of skp1 suppressor; subunit of SCF ubiquitin ligase complex; essential regulator of cell cycle; essential component of kinetochore assembly pathway.
97	YOL117W	RRI2	COP9 signalosome (CSN) subunit
98	NC-001146	NA	NA
99	YOL153C		Hypothetical ORF
100	YNR006W	VPS27	cysteine rich putative zinc finger essential for function| hydrophilic protein
101	YNR007C	ATG3	Protein involved in autophagocytosis during starvation
102	YNL194C		Integral membrane protein localized to eisosomes, large immobile protein structures at the cell cortex associated with endocytosis; sporulation and plasma membrane sphingolipid content are altered in mutants; has homologs SUR7 and FMP45
103	YNL230C	ELA1	elongin A transcription elongation factor
104	YNL325C	FIG4	Protein required for efficient mating, member of a family of eukaryotic proteins that contain a domain homologous to Sac1p
105	YMR316W	DIA1	may be involved in invasive growth, pseudohyphal growth
106	YMR291W		Hypothetical ORF
107	YMR196W		Hypothetical ORF
108	YMR173W	DDR48	contains >35 repeats of the amino acid sequence NNNDSYGS flocculent specific protein

109	YMR174C	PAI3	inhibitor of proteinase Pep4p
110	YMR175W	SIP18	Salt-Induced Protein
111	YMR169C	ALD3	aldehyde dehydrogenase
112	YMR139W	RIM11	Required for Ime1p phosphorylation, association of the Ime1p-Ume6p meiotic activator, early meiotic gene expression, and sporulation
113	YMR118C		Hypothetical ORF
114	YMR092C	AIP1	actin cortical patch component
115	YMR110C	HFD1	Hypothetical ORF
116	YMR087W		Putative ADP-ribose-1''-monophosphatase that converts ADP-ribose-1''-monophosphate to ADP-ribose; may have a role in tRNA splicing; contains an A1pp domain
117	YMR034C		Hypothetical ORF
118	YML002W		Hypothetical ORF
119	YML054C	CYB2	L-lactate cytochrome c oxidoreductase|cytochrome b2
120	YML100W	TSL1	similar to TPS3 gene product|trehalose-6-phosphate synthase/phosphatase complex 123 kDa regulatory subunit
121	YLR345W		Hypothetical ORF

5.3.4.3 Cluster 3

Cluster 3 contains 220 ORFs with upregulated expression in *glc7-12* mutant strains (Table 5.3). From the 220 ORFs 44 are uncharacterized and 59 are not annotated yet.

Interestingly, this cluster contains Sds22p and Reg2p, two known Glc7p regulatory subunits (Hong et al., 2000; MacKelvie et al., 1995), suggesting that the cell could try to rescue the absence of Glc7p by overexpressing subunits that target or regulate Glc7p in specific cellular processes. This also suggests that it is possible to identify regulatory subunits of Glc7p by our approach.

Intriguingly, we also find Yak1p and Crf1p in this cluster, two proteins implicated in the repression of RP genes (Martin et al., 2004). The location of Yak1p and Crf1p is redirected to the nucleus in *glc7-12* and *glc7-13* mutant cells, indicating that this might be the reason for the downregulation of RP genes in *glc7-12* mutants (see Chapter 6 and Chapter 5, Cluster 4).

In addition, Ino1p is listed in this cluster. Upregulation of Ino1p explains the INO resistant phenotype in *glc7-12* mutants and is probably also the reason for the resistance of the other *glc7* mutants.

Furthermore, we found Rec104p and Trs31p in cluster 3. The upregulation of these genes is due to their location in the genome downstream of snR71 and snR13. Transcription termination of these snoRNAs does not occur in *glc7-12*, thus tandem

transcripts (snoRNA-mRNA) are produced, leading to a higher number of the mRNA transcripts (see Chapter 3).

Finally, Hrp1p/Nab4p, a component of CF I is found in this cluster.

Table 5.3

no.	systematic name	gene name	description
1	YLR281C		Hypothetical ORF
2	YLR284C	ECI1	d3,d2-Enoyl-CoA Isomerase
3	YLR267W		bypass of PAM1
4	YLR193C	UPS1	Mitochondrial intermembrane space protein that regulates alternative processing and sorting of Mgm1p and other proteins;required for normal mitochondrial morphology; ortholog ofhuman PRELI
5	YLR194C		Hypothetical ORF, Structural constituent of the cell wall attached to the plasma membrane by a GPI-anchor; expression is upregulated in response to cell wall stress
6	YLR202C	NA	NA
7	YLR205C	HMX1	Homology to heme oxygenases
8	YLR206W	ENT2	Epsin-like protein required for endocytosis and actin patch assembly, functionally redundant with Ent1p; contains clathrin-binding motif at C terminus
9	YLR174W	IDP2	NADP-dependent isocitrate dehydrogenase
10	YLR177W		Hypothetical ORF
11	YLR149C		Hypothetical ORF
12	YLL020C	NA	NA
13	YLL020C	NA	NA
14	YLL019C	KNS1	protein kinase homolog
15	YLL057C		similar to Fe(II)-dependent sulfonate/alpha-ketoglutarate dioxygenase; EC 1.14.11.17
16	NC-001143	NA	NA
17	NC-001143	NA	NA
18	YKR091W	SRL3	Suppressor of Rad53 null Lethality
19	YKR096W		Hypothetical ORF
20	YKR064W	OAF3	Hypothetical ORF
21	YKR045C		Hypothetical ORF
22	YKR058W	GLG1	glycogen synthesis initiator
23	YKR014C	YPT52	rab5-like GTPase involved in vacuolar protein sorting and endocytosis
24	YKL086W		ATP-dependent cysteine sulfinic acid reductase
25	YKL124W	SSH4	Suppressor of SHR3; confers leflunomide resistance when overexpressed
26	YKL159C	RCN1	calcineurin inhibitor
27	YKL171W		Hypothetical ORF
28	YKL193C	SDS22	Glc7p regulatory subunit
29	YKL187C		Hypothetical ORF
30	YKL217W	JEN1	carboxylic acid transporter protein homolog
31	YJR149W		Hypothetical ORF
32	YJR091C	JSN1	benomyl dependent tubulin mutant
33	YJL142C	NA	NA
34	YJL142C	NA	NA

35	YJL141C	YAK1	Serine-threonine protein kinase
36	YJL153C	INO1	L-myo-inositol-1-phosphate synthase
37	NC-001133	NA	NA
38	YAR029W		Hypothetical ORF
39	YAR033W	MST28	Multicopy suppressor of Sec Twenty one
40	YAL017W	PSK1	contains serine/threonine protein kinase domain and shows homology with the SNF1 serine/threonine protein kinase
41	YAL051W	OAF1	transcription factor
42	NC-001148	NA	NA
43	NC-001148	NA	NA
44	NC-001147	NA	NA
45	YERWDELTA12	NA	NA
46	YDRCDELTA9	NA	NA
47	TV(UAC)B	NA	NA
48	TA(UGC)O	NA	NA
49	TG(UCC)N	NA	NA
50	YNLWSIGMA2	NA	NA
51	RDN37-1	NA	NA
52	RDN25-1	NA	NA
53	RDN5-1	NA	NA
54	TA(UGC)L	NA	NA
55	TL(UAA)J	NA	NA
56	TA(UGC)A	NA	NA
57	SUC4	NA	NA
58	NC-001141	NA	NA
59	NC-001141	NA	NA
60	YIL162W	SUC2	invertase (sucrose hydrolyzing enzyme)
61	YIR041W		Hypothetical ORF
62	YIL024C		Hypothetical ORF
63	YIL046W	MET30	contains five copies of WD40 motif and interacts with and regulates Met4p
64	YIL059C	NA	NA
65	YIL057C		Hypothetical ORF
66	YIL056W	VHR1	Hypothetical ORF, Transcriptional activator, required for the vitamin H-responsive element (VHRE) mediated induction of VHT1 (Vitamin H transporter) and BIO5 (biotin biosynthesis intermediate transporter) in response to low biotin concentrations
67	YIL050W	PCL7	cyclin
68	YIL077C		Hypothetical ORF
69	YIL108W		Hypothetical ORF
70	YIL105C	SLM1	Phosphoinositide PI4,5P(2) binding protein, forms a complex with Slm2p; acts downstream of Mss4p in a pathway regulating actin cytoskeleton organization in response to stress; subunit of and phosphorylated by the TORC2 complex, AVO2 interactor
71	YIL101C	XBP1	transcriptional repressor
72	YIL099W	SGA1	glucoamylase
73	YIL160C	POT1	3-oxoacyl CoA thiolase
74	NC-001140	NA	NA
75	NC-001140	NA	NA
76	YHR157W	REC104	meiosis-specific protein
77	YHR139C	SPS100	sporulation-specific cell wall maturation protein
78	YHR097C		Hypothetical ORF
79	YHL022C	SPO11	early meiosis-specific recombination protein
80	YHL021C		Mitochondrial protein of unknown function
81	YHL047C	ARN2	triacetylfusarinine C transporter
82	NC-001139	NA	NA
83	NC-001139	NA	NA

84	YGR236C	SPG1	Protein required for survival at high temperature during stationary phase; not required for growth on nonfermentable carbon sources
85	YGR243W	FMP43	Hypothetical ORF
86	YGR197C	SNG1	Involved in nitrosoguanidine resistance
87	YGR122W		Hypothetical ORF
88	YGR088W	CTT1	catalase T
89	YGR067C		Hypothetical ORF; has similarity to Adr1p DNA-binding domain
90	YGR052W	FMP48	Hypothetical ORF
91	YGL052W	NA	NA
92	YGL051W	MST27	protein with COPI and COPII bindng motifs
93	YGL046W	NA	NA
94	YGL045W	RIM8	Regulator of IME2 (RIM) Involved in proteolytic processing of Rim1p
95	YGL184C	STR3	cystathionine beta-lyase
96	YGL180W	ATG1	protein kinase
97	YGL247W	BRR6	nuclear envelope protein
98	NC-001138	NA	NA
99	NC-001138	NA	NA
100	NC-001138	NA	NA
101	YFR022W	ROG3	Protein that binds to Rsp5p, which is a hect-type ubiquitin ligase, via its 2 PY motifs; has similarity to Rod1p; mutation suppresses the temperature sensitivity of an mck1 rim11 double mutant
102	RPR1	NA	NA
103	NC-001137	NA	NA
104	NC-001137	NA	NA
105	NC-001137	NA	NA
106	YER179W	DMC1	meiosis-specific protein related to RecA and Rad51p. Dmc1p colocalizes with Rad51p to discrete subnuclear sites in nuclear spreads during mid prophase, briefly colocalizes with Zip1p, and then disappears by pachytene
107	YER144C	UBP5	ubiquitin-specific protease (putative)
108	YER067W		Hypothetical ORF
109	YER020W	GPA2	nucleotide binding regulatory protein
110	NC-001136	NA	NA
111	NC-001136	NA	NA
112	NC-001136	NA	NA
113	NC-001136	NA	NA
114	NC-001136	NA	NA
115	NC-001136	NA	NA
116	YDR525W-A	SNA2	Homology to PMP3/SNA1 (Sensitivity to Na+)
117	YDR472W	TRS31	targeting complex (TRAPP) component involved in ER to Golgi membrane traffic
118	YDR402C	DIT2	catalyzes oxidation of N-formyl tyrosine to N,N-bisformyl dityrosine *in vitro*|cytochrome P450 56
119	YDR247W	VHS1	Gene whose overexpression suppresses the synthetic lethality of the hal3 sit4 double mutation
120	YDR253C	MET32	highly homologous to Met31p|transcriptional regulator of sulfur amino acid metabolism|zinc finger protein
121	YDR256C	CTA1	catalase A
122	YDR223W	CRF1	Transcriptional corepressor involved in the regulation of ribosomal protein gene transcription via the TOR signaling pathway and protein kinase A, phosphorylated by activated Yak1p which promotes accumulation of Crf1p in the nucleus
123	YDR213W	UPC2	zinc finger transcription factor of the Zn(2)-Cys(6) binuclear cluster domain type
124	YDR129C	SAC6	actin filament bundling protein|fimbrin homolog
125	YDR044W	HEM13	coproporphyrinogen III oxidase
126	YDR042C		Hypothetical ORF
127	YDR006C	SOK1	gene dosage suppressors of the conditional growth defect of several temperature-sensitive A kinase mutants
128	YDL010W		Hypothetical ORF

129	YDL008W	APC11	anaphase promoting complex (APC) subunit
130	YDL059C	RAD59	the RAD59 gene product has homology to the Rad52 protein
131	YDL079C	MRK1	MDS1 related protein kinase
132	YDL215C	GDH2	NAD-dependent glutamate dehydrogenase
133	YDL234C	GYP7	GTPase activating protein (GAP)
134	YCR101C		Hypothetical ORF
135	YCL039W	GID7	Involved in proteasome-dependent catabolite inactivation of fructose-1,6-bisphosphatase
136	YCL027W	FUS1	cell-surface protein required for cell fusion
137	YCL048W	SPS22	Protein of unknown function, redundant with Sps2p for the organization of the beta-glucan layer of the spore wall
138	NC-001134	NA	NA
139	YBR109w-a	NA	NA
140	YBR284W		Hypothetical ORF
141	YBR298C	MAL31	maltose permease
142	YBR280C	SAF1	F-Box protein involved in proteasome-dependent degradation of Aah1p during entry of cells into quiescence; interacts with Skp1
143	YBR150C	TBS1	Probable Zn-finger protein
144	YBR099C	NA	NA
145	YBR071W		Hypothetical ORF
146	YBR050C	REG2	Glc7p regulatory subunit
147	YBR016W		Hypothetical ORF
148	YBR005W	RCR1	Protein of the endoplasmic reticulum membrane involved in chitin deposition in the cell wall; overproduction confers resistance to Congo Red
149	NC-001148	NA	NA
150	YPR152C	URN1	Hypothetical ORF, Pre-mRNA splicing factor that is associated with the U2-U5-U6 snRNPs, the RES complex, and the Prp19-associated complex (NTC); null mutation displays synthetic genetic interactions with mutations affecting U2 snRNA and pre-mRNA splicing factors
151	YPR038W	NA	NA
152	YPR039W	NA	NA
153	YPR002W	PDH1	prpD homologue; (62% identical to the prpD genes of Escherichia coli and Salmonella typhimurium, which play an unknown but essential role in propionate catabolism)
154	YPL054W	LEE1	Product of gene unknown
155	YPL100W	ATG21	Maturation of pro-AmInopeptidase I (proAPI) defective; protein similar to Aut10p and YGR223Cp with three putative WD repeats
156	YPL201C	YIG1	Protein that interacts with glycerol 3-phosphatase and plays a role in anaerobic glycerol production; localizes to the nucleus and cytosol
157	YPL200W	CSM4	Protein required for accurate chromosome segregation during meiosis
158	YPL230W		Up in StarVation
159	YPL247C		Hypothetical ORF
160	NC-001147	NA	NA
161	NC-001147	NA	NA
162	NC-001147	NA	NA
163	NC-001147	NA	NA
164	YOR385W		Hypothetical ORF
165	NC-001147	NA	NA
166	YOR374W	ALD4	aldehyde dehydrogenase
167	YOR343C	NA	NA
168	YOR328W	PDR10	ABC transporter (putative)|highly similar to Pdr5p
169	YOR278W	HEM4	uroporphyrinogen III synthase
170	YOR219C	STE13	dipeptidyl aminopeptidase
171	YOR220W	WSP1	Hypothetical ORF, protein that interacts genetically with the Pat1 mRNA-decapping factor
172	YOR186W		Hypothetical ORF

173	YOR134W	BAG7	GTPase activating protein (GAP)
174	YOR050C	NA	NA
175	YOR018W	ROD1	Membrane protein; overexpression confers resistance to the GST substrate o-dinitrobenzene as well as to zinc and calcium; contains a PY-motif, which is required for Rod1p interaction with Rsp5p, a hect-type ubiquitin ligase
176	YOR019W		Hypothetical ORF
177	YOR023C	AHC1	Ada histone acetyltransferase complex component
178	YOL015W	IRC10	Hypothetical ORF
179	YOL016C	CMK2	calmodulin-dependent protein kinase
180	YOL089C	HAL9	contains zinc finger|transcription factor (putative)
181	YOL126C	MDH2	malate dehydrogenase
182	YOL123W	HRP1	cleavage and polyadenylation factor CF I component involved in pre-mRNA 3' end processing
183	YOL113W	SKM1	Serine/threonine protein kinase with similarity to Ste20p and Cla4p
184	YOL162W		Hypothetical ORF, member of the Dal5p subfamily of the major facilitator family
185	YOL154W	ZPS1	Zinc and pH regulated Surface protein. Similar to Candida albicans protein Pra1, a pH-regulated cell surface glycoprotein with weak similarity to zinc metalloproteinases.
186	YNR073C		Hypothetical ORF
187	YNR001C	CIT1	citrate synthase
188	YNL020C	ARK1	serine/threonine kinase (putative)
189	YNL008C	ASI3	Amino acid Sensor-Independent (ASI) genes encode membrane proteins Asi1p, Asi2p and Asi3p. Asi1p and Asi3p have conserved ubiquitin ligase-like RING domains at their C-termini
190	YNL077W	APJ1	J-protein co-chaperone family 20 kDa
191	YNL120C	NA	NA
192	YNL207W	RIO2	Protein required for cell viability
193	YNL195C		Hypothetical ORF
194	YNL192W	CHS1	chitin synthase 1
195	NC-001145	NA	NA
196	NC-001145	NA	NA
197	NC-001145	NA	NA
198	NC-001145	NA	NA
199	YMR280C	CAT8	zinc-cluster protein involved in activating gluconeogenic genes; related to Gal4p
200	YMR206W		Hypothetical ORF
201	YMR152W	YIM1	protease|similar to E. coli leader peptidase
202	YMR161W	HLJ1	Homologous to E coli dnaJ protein
203	YMR135C	GID8	Dose-dependent Cell cycle Regulator
204	YMR136W	GAT2	Similar to GATA-family of DNA binding proteins
205	YMR096W	SNZ1	highly conserved 35 kDa protein that shows increased expression after entry into stationary phase
206	YMR101C	SRT1	cis-prenyltransferase
207	YMR102C		Hypothetical ORF
208	YMR103C	NA	NA
209	YMR107W	SPG4	Protein required for survival at high temperature during stationary phase; not required for growth on nonfermentable carbon sources
210	YMR081C	ISF1	May regulate NAM7 function, possibly at level of mRNA turnover
211	YMR084W		Hypothetical ORF
212	YMR085W		Hypothetical ORF
213	YML014W		mcm5U/mcm5s2U tRNA carboxyl methyltransferase
214	NC-001144	NA	NA
215	NC-001144	NA	NA
216	NC-001144	NA	NA
217	YLR414C		Hypothetical ORF
218	YLR350W	ORM2	Endoplasmic reticulum membrane-anchored protein

219	YLR327C	TMA10	Protein of unknown function that associates with ribosomes
220	Z73326	NA	NA

5.3.4.4 Cluster 4

This cluster contains 182 ORFs the expression of which is downregulated in *glc7-12* mutant strains (Table 5.4). From the 182 ORFs 9 are uncharacterized ORFs and 8 are not annotated yet.

Most striking about cluster 4 is that it contains a large amount of RP (34) and Ribi genes (32; see Table 6.1, Chapter 6), which corresponds to ~36% of the genes in this cluster. RP and Ribi genes function in ribosome biogenesis. Because these two groups of genes are found in the same cluster, it is likely that they have a common transcriptional response to changes in the environment, suggesting that their expression is controlled by the same transcription factor(s). A transcription factor known to regulate the Ribi regulon and the RP regulon is Sfp1p (Jorgensen et al., 2004). Therefore, it is tempting to speculate that Glc7p might regulate Sfp1p to coordinate transcription of RP and Ribi genes.

Table 5.4

no.	systematic name	gene name	description
1	YLR293C	GSP1	GTP-binding protein
2	YLR249W	YEF3	Translation elongation factor 3 (EF-3)
3	YLR195C	NMT1	N-myristoyl transferase
4	YLR186W	EMG1	ribosome biogenesis
5	YLR188W	MDL1	ATP-binding cassette (ABC) transporter family member
6	YLR129W	DIP2	U3 snoRNP protein
7	YLR075W	RPL10	ribosomal protein L10
8	YLR061W	RPL22A	ribosomal protein L22A (L1c) (rp4) (YL31)
9	YLR048W	RPS0B	ribosomal protein S0B
10	YLR048W	RPS0B	ribosomal protein S0B
11	YLR017W	MEU1	Protein that regulates ADH2 gene expression
12	YLL045C	RPL8B	ribosomal protein L8B (L4B) (rp6) (YL5)
13	YKR043C		Hypothetical ORF
14	YKR026C	GCN3	eIF2B 34 kDa alpha subunit
15	YKL056C	TMA19	Protein of unknown function that associates with ribosomes; homolog of translationally controlled tumor protein; green fluorescent protein (GFP)-fusion protein localizes to the cytoplasm; YKL056C is not an essential gene
16	YKL081W	TEF4	translation elongation factor EF-1gamma

17	YKL113C	RAD27	42 kDa 5' to 3' exonuclease required for Okazaki fragment processing
18	YKL130C	SHE2	Required for mother cell-specific HO expression
19	YKL152C	GPM1	phosphoglycerate mutase
20	YKL181W	PRS1	ribose-phosphate pyrophosphokinase
21	YKL180W	RPL17A	ribosomal protein L17A (L20A) (YL17)
22	YKL180W	RPL17A	ribosomal protein L17A (L20A) (YL17)
23	YJR123W	RPS5	ribosomal protein S5 (S2) (rp14) (YS8)
24	YJR069C	HAM1	Product of gene unknown
25	YJR041C	URB2	Nucleolar protein required for normal metabolism of the rRNA primary transcript, proposed to be involved in ribosome biogenesis; Protein required for cell viability
26	YJR031C	GEA1	GDP/GTP exchange factor
27	YJR007W	SUI2	Translation initiation factor eIF-2 alpha subunit
28	YJL111W	CCT7	chaperonin containing T-complex subunit seven component
29	YJL109C	UTP10	U3 snoRNP protein
30	YJL104W	PAM16	Constituent of the mitochondrial import motor associated with the presequence translocase, along with Ssc1p, Tim44p, Mge1p, and Pam18p; has similarity to J-domain containing proteins mitochondrial protein of unknown function
31	YJL138C	TIF2	translation initiation factor eIF4A subunit
32	YJL177W	RPL17B	ribosomal protein L17B (L20B) (YL17)
33	YJL177W	RPL17B	ribosomal protein L17B (L20B) (YL17)
34	YJL200C	ACO2	Hypothetical ORF, Putative mitochondrial aconitase isozyme; similarity to Aco1p, an aconitase required for the TCA cycle; expression induced during growth on glucose, by amino acid starvation via Gcn4p, and repressed on ethano
35	YJL198W	PHO90	Low-affinity phosphate transporter
36	YJL190C	RPS22A	ribosomal protein S22A (S24A) (rp50) (YS22)
37	NC-001139	NA	NA
38	NC-001139	NA	NA
39	NC-001134	NA	NA
40	YIL043C	CBR1	cytochrome b reductase
41	YIL078W	THS1	threonine-tRNA ligase
42	YHR208W	BAT1	branched-chain amino acid transaminase|highly similar to mammalian ECA39, which is regulated by the oncogene myc
43	YHR108W	GGA2	ARF-binding protein
44	YHR128W	FUR1	UPRTase
45	YHR089C	GAR1	small nucleolar RNP protein
46	YHR092C	HXT4	high affinity glucose transporter
47	YHR072W	ERG7	2,3-oxidosqualene-lanosterol cyclase
48	YHR020W		Protein required for cell viability
49	YHR024C	MAS2	mitochondrial processing protease 53 kDa subunit
50	YGR214W	RPS0A	ribosomal protein S0A
51	YGR214W	RPS0A	ribosomal protein S0A
52	YGR159C	NSR1	nuclear localization sequence binding protein
53	YGR123C	PPT1	serine/threonine phosphatase
54	YGR082W	TOM20	20 kDa mitochondrial outer membrane protein import receptor
55	YGL008C	PMA1	plasma membrane H+-ATPase
56	YGL076C	RPL7A	ribosomal protein L7A (L6A) (rp11) (YL8)
57	YGL123W	RPS2	ribosomal protein S2 (S4) (rp12) (YS5)
58	YGL148W	ARO2	chorismate synthase
59	YGL143C	MRF1	mitochondrial polypeptide chain release factor
60	YFR031C-A	RPL2A	ribosomal protein L2A (L5A) (rp8) (YL6)
61	YFR009W	GCN20	ATP-binding cassette (ABC) family
62	YFL022C	FRS2	phenylalanine-tRNA ligase subunit
63	YFL045C	SEC53	phosphomannomutase
64	YER156C		Hypothetical ORF

65	YER055C	HIS1	ATP phosphoribosyltransferase
66	YER056C	FCY2	purine-cytosine permease
67	YER023W	PRO3	delta 1-pyrroline-5-carboxylate reductase
68	YER025W	GCD11	translational initiation factor eIF-2 gamma subunit
69	YER003C	PMI40	mannose-6-phosphate isomerase
70	YEL001C		Hypothetical ORF
71	YEL040W	UTR2	(alias: Congo Red Hypersensitivity)
72	YDR492W	IZH1	Membrane protein involved in zinc metabolism, member of the four-protein IZH family; transcription is regulated directly by Zap1p, expression induced by zinc deficiency and fatty acids; deletion increases sensitivity to elevated zinc
73	YDR487C	RIB3	3,4-dihydroxy-2-butanone 4-phosphate synthase
74	YDR454C	GUK1	guanylate kinase
75	YDR429C	TIF35	translation initiation factor eIF3 subunit
76	YDR399W	HPT1	hypoxanthine guanine phosphoribosyltransferase
77	YDR418W	RPL12B	ribosomal protein L12B (L15B) (YL23)
78	YDR382W	RPP2B	ribosomal protein P2B (YP2beta) (L45)
79	YDR361C	BCP1	Hypothetical ORF
80	YDR341C	RRS1	Cytoplasmic arginyl-tRNA synthetase; Protein required for cell viability
81	YDR260C	SWM1	Spore Wall Maturation 1
82	YDR237W	MRPL7	Mitochondrial ribosomal protein of the large subunit
83	YDR174W	HMO1	high mobility group (HMG) family
84	YDR091C		ATP-binding cassette (ABC) superfamily nontransporter group (putative)
85	YDR064W	RPS13	ribosomal protein S13 (S27a) (YS15)
86	YDR033W	MRH1	Membrane protein related to Hsp30p; Localized by immunofluorescence to membranes, mainly the plasma membr. punctuate immunofluorescence pattern observed in buds. The nuclear envelope, but not vacuole or mitochondrial membranes also stained
87	YDL081C	RPP1A	acidic ribosomal protein P1A (YP1alpha) (A1)
88	YDL112W	TRM3	tRNA (Gm18) ribose methylase
89	YDL130W	RPP1B	ribosomal protein P1B (L44') (YP1beta) (Ax)
90	YDL130W	RPP1B	ribosomal protein P1B (L44') (YP1beta) (Ax)
91	YDL155W	CLB3	B-type cyclin
92	YDL158C	NA	NA
93	YDL192W	ARF1	ADP-ribosylation factor
94	YCR034W	FEN1	Fatty acid elongase, involved in sphingolipid biosynthesis; acts on fatty acids of up to 24 carbons in length; mutations have regulatory effects on 1,3-beta-glucan synthase, vacuolar ATPase, and the secretory pathway
95	YBR249C	ARO4	3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthase isoenzyme
96	YBR196C	PGI1	glucose-6-phosphate isomerase
97	YBR189W	RPS9B	ribosomal protein S9B (S13) (rp21) (YS11)
98	YBR127C	VMA2	vacuolar ATPase V1 domain subunit B (60 kDa)
99	YBR133C	HSL7	Has homology to arginine methyltransferases
100	YBR143C	SUP45	eRF1 (eukaryotic Release Factor 1) homolog
101	YBR146W	MRPS9	ribosomal protein S9 (putative)
102	YBR104W	YMC2	Mitochondrial carrier protein
103	YBR093C	PHO5	acid phosphatase
104	YBR069C	TAT1	amino acid transport protein for valine, leucine, isoleucine, and tyrosine
105	YBR078W	ECM33	ExtraCellular Mutant
106	YBR043C	QDR3	Multidrug transporter required for resistance to quinidine, barban, cisplatin, and bleomycin; member of the major facilitator superfamily of transporters conferring multiple drug resistance (MFS-MDR)
107	YBR048W	RPS11B	ribosomal protein S11B (S18B) (rp41B) (YS12)

108	YBR025C		Hypothetical ORF
109	YBR031W	RPL4A	ribosomal protein L4A (L2A) (rp2) (YL2)
110	YBL004W	UTP20	Component of the small-subunit (SSU) processome, which is involved in the biogenesis of the 18S rRNA; U3 snoRNP protein
111	YPR114W		Hypothetical ORF
112	YPR074C	TKL1	transketolase 1
113	YPR033C	HTS1	histidine-tRNA ligase
114	YPR010C	RPA135	RNA polymerase I subunit
115	YPL019C	VTC3	polyphosphate synthetase (putative)
116	YPL037C	EGD1	pol II transcribed genes regulator
117	YPL086C	ELP3	histone acetyltransferase|RNA polymerase II Elongator subunit
118	YPL131W	RPL5	ribosomal protein L5 (L1a)(YL3)
119	YPL211W	NIP7	Nip7p is required for 60S ribosome subunit biogenesis
120	YPL198W	RPL7B	ribosomal protein L7B (L6B) (rp11) (YL8)
121	YPL226W	NEW1	This gene encodes a protein with an Q/N-rich amino terminal domain that acts as a prion, termed [NU]+.
122	YPL220W	RPL1A	ribosomal protein L1A, forms part of the 60S ribosomal subunit
123	YPL256C	CLN2	G1 cyclin
124	YPL273W	SAM4	AdoMet-homocysteine methyltransferase
125	YOR341W	RPA190	RNA polymerase I subunit
126	YOR323C	PRO2	gamma-glutamyl phosphate reductase
127	YOR335C	ALA1	Cytoplasmic alanyl-tRNA synthetase gene
128	YOR272W	YTM1	microtubule-associated protein
129	YOR243C	PUS7	pseudouridine synthase
130	YOR247W	SRL1	Suppressor of Rad53 null Lethality
131	YOR248W	NA	NA
132	YOR229W	WTM2	transcriptional modulator
133	YOR063W	RPL3	ribosomal protein L3 (rp1) (YL1)
134	YOL040C	RPS15	ribosomal protein S15 (S21) (rp52) (RIG protein)
135	YOL039W	RPP2A	60S acidic ribosomal protein P2A (L44) (A2) (YP2alpha)
136	YOL077C	BRX1	Essential nucleolar protein required for biogenesis of the 60S ribosomal subunit
137	YOL112W	MSB4	Multicopy Suppressor of Bud Emergence
138	NC-001146	NA	NA
139	YNR050C	LYS9	Seventh step in lysine biosynthesis pathway
140	YNL112W	DBP2	ATP dependent RNA helicase|dead box protein
141	YNL137C	NAM9	mitochondrial S4 ribosomal protein (putative)
142	YNL135C	FPR1	peptidyl-prolyl cis-trans isomerase (PPIase)
143	YNL123W	NMA111	Protein of unknown function which may contribute to lipid homeostasis and/or apoptosis; sequence similarity to the mammalian Omi/HtrA2 family of serine proteases
144	YNL178W	RPS3	ribosomal protein S3 (rp13) (YS3)
145	YNL218W	MGS1	Maintenance of Genome Stability 1
146	YNL209W	SSB2	SSB1 homolog|heat shock protein of HSP70 family
147	YNL256W	FOL1	dihydro-6-hydroxymethylpterin pyrophosphokinase| dihydroneopterin aldolase|dihydropteroate synthetase
148	YNL255C	GIS2	GIG3 suppressor
149	YNL302C	RPS19B	ribosomal protein S19B (rp55B) (S16aB) (YS16B)
150	YNL301C	RPL18B	ribosomal protein L18B (rp28B)
151	SNR72	NA	NA
152	NC-001145	NA	NA
153	YMR321C		Hypothetical ORF
154	YMR307W	GAS1	cell surface glycoprotein 115-120 kDa
155	YMR308C	PSE1	karyopherin
156	YMR314W	PRE5	20S proteasome alpha-type subunit
157	YMR290C	HAS1	RNA-dependent helicase (putative)

158	YMR241W	YHM2	DNA binding protein|mtDNA stabilizing protein, mitochondrial inner membrane protein with low homology to RIM2
159	YMR217W	GUA1	GMP synthase
160	YMR202W	ERG2	C-8 sterol isomerase
161	YMR146C	TIF34	translation initiation factor eIF3 p39 subunit
162	YMR116C	ASC1	G-beta like protein
163	YMR116C	ASC1	G-beta like protein
164	YMR129W	POM152	membrane glycoprotein|nuclear pore complex subunit
165	YMR079W	SEC14	phosphatidylinositol transfer protein
166	YMR049C	ERB1	Protein required for maturation of the 25S and 5.8S ribosomal RNAs; homologous to mammalian Bop1
167	YMR033W	ARP9	actin related protein|chromatin remodeling Snf/Swi complex subunit
168	YML022W	APT1	adenine phosphoribosyltransferase
169	YML056C	IMD4	IMP dehydrogenase homolog
170	YML078W	CPR3	cyclophilin|peptidyl-prolyl cis-trans isomerase (PPIase)
171	YML106W	URA5	orotate phosphoribosyltransferase 1
172	YML123C	PHO84	inorganic phosphate transporter
173	YML116W	ATR1	very hydrophobic, has many membrane-spanning regions, several potential glycosylation sites, potential ATP-binding site
174	YLR448W	RPL6B	ribosomal protein L6B (L17B) (rp18) (YL16)
175	YLR409C	UTP21	Possible U3 snoRNP protein involved in maturation of pre-18S rRNA, based on computational analysis of large-scale protein-protein interaction data; U3 snoRNP protein
176	YLR413W		Hypothetical ORF
177	YLR418C	CDC73	accessory factor associated with RNA polymerase II by affinity chromatography
178	YLR384C	IKI3	RNA polymerase II Elongator subunit
179	YLR367W	RPS22B	ribosomal protein S22B (S24B) (rp50) (YS22)
180	YLR372W	SUR4	elongase
181	YLR359W	ADE13	adenylosuccinate lyase
182	YLR325C	RPL38	ribosomal protein L38

5.3.4.5 Cluster 5

Cluster 5 contains 76 ORFs with upregulated expression in *glc7-12* mutant strains (Table 5.5). From the 76 ORFs, 3 are uncharacterized ORFs and 28 are not annotated yet.

This cluster contains Nrd1p, a component of the Nrd1 complex that is involved in snoRNA transcription termination (Steinmetz et al., 2001). A consequence of mutations in subunits of the Nrd1 complex is that in addition to the snoRNA transcription termination defect the Nrd1 message itself accumulates. The upregulation of Nrd1 mRNA levels observed in this microarray experiment was verified by Northern analysis (see Chapter 3). Strong accumulation of Nrd1 was observed in several glc7 mutants. Therefore, two independent experiments implicate Glc7p in the transcriptional autoregulation mechanism of Nrd1p.

Table 5.5

no.	systematic name	gene name	description
1	YLR256W	HAP1	zinc finger transcription factor of the Zn(2)-Cys(6) binuclear cluster domain type
2	YLR134W	PDC5	pyruvate decarboxylase
3	YLR136C	TIS11	Zinc finger containing homolog of mammalian TIS11, glucose repressible gene
4	YLR099C	ICT1	Increased Copper Tolerance; Similar to Ecm18p
5	YLR042C		Hypothetical ORF
6	NC-001143	NA	NA
7	NC-001143	NA	NA
8	YKR069W	MET1	Methionine metabolism
9	YKR052C	MRS4	carrier protein highly homologous to Mrs3p
10	YKR011C		Hypothetical ORF
11	YKL044W	NA	NA
12	YKL116C	PRR1	protein kinase
13	YKL164C	PIR1	contains tandem internal repeats
14	NC-001133	NA	NA
15	NC-001133	NA	NA
16	YAL001C	TFC3	138 kDa|transcription factor tau (TFIIIC) subunit
17	YAL016W	TPD3	protein phosphatase 2A regulatory subunit A
18	YAL040C	CLN3	G1 cyclin
19	NC-001137	NA	NA
20	NC-001144	NA	NA
21	NC-001145	NA	NA
22	TG(GCC)D2	NA	NA
23	TG(CCC)D	NA	NA
24	TP(AGG)C	NA	NA
25	TG(GCC)C	NA	NA
26	TG(GCC)B	NA	NA
27	TS(UGA)P	NA	NA
28	TG(GCC)O1	NA	NA
29	TP(AGG)N	NA	NA
30	RDN5-1	NA	NA
31	RDN37-1	NA	NA
32	TH(GUG)K	NA	NA
33	TR(ACG)J	NA	NA
34	YIR017C	MET28	transcriptional activator in the Cbf1p-Met4p-Met28p complex
35	YIL109C	SEC24	vesicle coat component
36	YHR029C	YHI9	Protein of unknown function that is a member of the PhzF superfamily, although unlike its bacterial homolog, is most likely not involved in phenazine production; possibly involved in a membrane regulation metabolic pathway
37	YHL036W	MUP3	very low affinity methionine permease
38	YGR055W	MUP1	high affinity methionine permease
39	YGL062W	PYC1	pyruvate carboxylase
40	YGL125W	MET13	methylenetetrahydrofolate reductase (mthfr) (putative)
41	YFR015C	GSY1	glycogen synthase (UDP-glucose-starch glucosyltransferase)
42	YER091C	MET6	vitamin B12-(cobalamin)-independent isozyme of methionine synthase (also called N5-methyltetrahydrofolate homocysteine methyltransferase or 5-methyltetrahydropteroyl triglutamate homocysteine methyltransferase)
43	YER093C-A		Hypothetical ORF

44	YEL071W	DLD3	D-lactate dehydrogenase
45	NC-001136	NA	NA
46	YDR516C	EMI2	Early Meiotic Induction
47	YDR488C	PAC11	Protein required in the absence of Cin8p
48	YDR380W	ARO10	Phenylpyruvate decarboxylase, catalyzes decarboxylation of phenylpyruvate to phenylacetaldehyde, which is the first specific step in the Ehrlich pathway
49	YDR381W	YRA1	RNA-binding RNA annealing protein
50	YDR207C	UME6	C6 zinc finger URS1-binding protein
51	YDR046C	BAP3	valine transporter
52	YDR011W	SNQ2	ABC transporter
53	YDL022W	GPD1	glycerol-3-phosphate dehydrogenase
54	YCR030C	SYP1	Suppressor of Yeast Profilin deletion
55	YCR005C	CIT2	citrate synthase
56	NC-001134	NA	NA
57	YBR212W	NGR1	glucose-repressible RNA binding protein
58	YBR179C	FZO1	Drosophila melanogaster fuzzy onions gene homolog integral protein of the mitochondrial outer membrane; can be isolated as part of a high molecular weight complex
59	YBL045C	COR1	coenzyme QH2 cytochrome c reductase 44 kDa core protein subunit
60	NC-001148	NA	NA
61	YPR038W	NA	NA
62	YPL192C	PRM3	pheromone-regulated membrane protein
63	YOR264W	DSE3	Daughter Specific Expression 3
64	YOR176W	HEM15	ferrochelatase (protoheme ferrolyase)
65	NC-001146	NA	NA
66	YNL078W	NIS1	
67	YNL212W	VID27	Vacuole import and degradation
68	YNL251C	NRD1	RNA-binding protein that interacts with the C-terminal domain of the RNA polymerase II large subunit (Rpo21p), required for transcription termination and 3' end maturation of nonpolyadenylated RNAs
69	YNL241C	ZWF1	glucose-6-phosphate dehydrogenase
70	YNL277W	MET2	homoserine O-trans-acetylase
71	NC-001145	NA	NA
72	NC-001145	NA	NA
73	YMR301C	ATM1	ABC transporter
74	YMR141C	NA	NA
75	YMR004W	MVP1	Protein required for sorting proteins to the vacuole
76	YML092C	PRE8	proteasome component Y7

5.3.4.6 Cluster 6

Cluster 6 contains 68 ORFs with downregulated expression in *glc7-12* mutant strains (Table 5.6). From the 68 ORFs 14 are uncharacterized and six are not annotated yet.
In addition to two known genes of the Ribi regulon, RNT1 and PUS6, cluster 6 also contains RPC25, a RNAP III subunit. Several subunits of RNAP I and III are accredited to the Ribi regulon since they transcribe distinct classes of rRNA, important components of ribosomes. We also identified different RNAP I subunits

(RPA190, RPA135, RPA49) and one RNAP III subunit (RPC34) in cluster 1 and 4 in our microarray experiment. It is most likely that RPC25 also belongs to the Ribi regulon but has not yet been identified as a member of this regulon. Interestingly, Rpa190p, the largest subunit of RNAP I contains a canonical RVXF motif, found in the majority of protein phosphatase type I interacting proteins (Walsh et al., 2002). In addition, Rpa190p is known to be a phospho-protein (Bell et al., 1976). Both proteins, Glc7p and Rpa190p, are enriched in the nucleolus. Taken together, it is highly likely that Rpa190p might be a Glc7p substrate.

RNT1 encodes an endonuclease involved in snoRNA transcription termination. Downregulation of RNT1 in glc7-12 cells could contribute or even be the cause of the snoRNA transcription termination defect observed in this mutant (see Chapter 3).

Moreover, cluster 6 contains a component of the TRAMP complex, TRF5 that is involved in polyadenylation-mediated RNA surveillance (Vanacova et al., 2005). It was shown that this complex plays a part in processing of snoRNAs and therefore downregulation of TRF5 might impair this function. This could also add to the snoRNA transcription termination defect observed in glc7 mutants.

Table 5.6

no.	systematic name	gene name	description
1	YLR237W	THI7	thiamine transporter
2	YLR244C	MAP1	methionine aminopeptidase
3	YLR018C	POM34	integral membrane protein|nuclear pore complex subunit
4	YLR021W		Hypothetical ORF
5	YKR063C	LAS1	May regulate expression of genes involved in bud formation and morphogenesis
6	YKL069W		Hypothetical ORF
7	YKL144C	RPC25	RNA polymerase III subunit
8	NC-001142	NA	NA
9	YJR097W	JJJ3	Protein of unknown function, contains a J-domain, which is a region with homology to the E. coli DnaJ protein
10	YJL073W	JEM1	DnaJ-like protein of the endoplasmic reticulum membrane
11	YDRWDELTA10	NA	NA
12	NC-001141	NA	NA
13	YIR032C	DAL3	ureidoglycolate hydrolase
14	NC-001140	NA	NA
15	YHR144C	DCD1	dCMP deaminase
16	YHR057C	CPR2	cyclophilin|peptidyl-prolyl cis-trans isomerase (PPIase)
17	YHL009C	YAP3	bZIP protein; transcription factor
18	NC-001139	NA	NA

19	YGR275W	RTT102	Regulator of Ty1 Transposition
20	YGR251W		Protein required for cell viability
21	YGR169C	PUS6	RNA:Psi-synthase
22	YGR125W		Hypothetical ORF
23	YGR079W		Hypothetical ORF
24	YGR021W		Hypothetical ORF
25	YGL079W		Hypothetical ORF
26	YGL157W		Oxidoreductase, catalyzes NADPH-dependent reduction of the bicyclic diketone bicyclo[2.2.2]octane-2,6-dione (BCO2,6D) to the chiral ketoalcohol (1R,4S,6S)-6-hydroxybicyclo[2.2.2]octane-2-one (BCO2one6ol)
27	YGL221C	NIF3	similar to Listeria monocytogenes major sigma factor (rpoD gene product)
28	YGL220W		Hypothetical ORF
29	YER163C		Hypothetical ORF
30	NC-001136	NA	NA
31	YEL066W	HPA3	Histone and other Protein Acetyltransferase; Has sequence homology to known HATs and NATs
32	YDR506C		Hypothetical ORF
33	YDR415C		Hypothetical ORF
34	YDR291W	HRQ1	Hypothetical ORF, Putative DNA helicase
35	YDR179W-A		Hypothetical ORF
36	YDR137W	RGP1	Ric1p-Rgp1p is an exchange factor, and peripheral membrane protein complex restricted to the Golgi.
37	YDR140W		Putative S-adenosylmethionine-dependent methyltransferase of the seven beta-strand family
38	YDR075W	PPH3	protein phosphatase type 2A
39	YDR034C	LYS14	Transcriptional activator of lysine pathway genes with 2-aminoadipate semialdehyde as co-inducer; saccharopine reductase synthesis
40	YDL219W	DTD1	D-Tyr-tRNA(Tyr) deacylase
41	YDL219W	DTD1	D-Tyr-tRNA(Tyr) deacylase
42	YCR060W	TAH1	HSP90 cofactor; interacts with Hsp82p, Pih1p, Rvb1 and Rvb2, contains a single TPR domain with at least two TPR motifs
43	YCL036W		Great for FULL DEAD box protein activity
44	YBR176W	ECM31	ExtraCellular Mutant
45	YBR188C	NTC20	splicing factor
46	YBR157C	ICS2	Increased Copper Sensitivity
47	YBR054W	YRO2	Homolog to HSP30 heat shock protein Yro1p
48	YPR162C	ORC4	origin recognition complex (ORC) 56 kDa subunit
49	YPR083W	MDM36	Mitochondrial Distribution and Morphology
50	YPL115C	BEM3	rho GTPase activating protein (GAP)
51	YPL214C	THI6	TMP pyrophosphorylase|hydroxyethylthiazole kinase
52	YOR337W	TEA1	Mutants are defective in Ty1 Enhancer-mediated Activation
53	YOR302W		CPA1 uORF , Arginine attenuator peptide, regulates translation of the CPA1 mRNA
54	YOR303W	CPA1	arginine specific|carbamoyl phosphate synthetase
55	YOR144C	ELG1	Enhanced Level of Genomic instability, Repressor of Ty1 Transposition
56	YOL026C		Protein required for cell viability
57	YNL023C	FAP1	transcription factor homolog; similarity to Drosophila melanogaster shuttle craft protein; similarity to human NFX1 protein; similarity to human DNA-binding protein tenascin
58	YNL022C		Non-essential protein with similarity to S. pombe hypothetical protein E349594
59	YNL065W	AQR1	multidrug resistance transporter
60	YNL299W	TRF5	DNA polymerase sigma
61	YMR319C	FET4	low affinity Fe2+ transport protein
62	YMR264W	CUE1	Ubc7p binding and recruitment protein

63	YMR236W	TAF9	TFIID subunit
64	YMR239C	RNT1	ribonuclease III
65	YMR226C		NADP(+)-dependent dehydrogenase; acts on serine, L-allo-threonine, and other 3-hydroxy acids
66	YMR074C		Hypothetical ORF
67	YLR376C		Platinum Sensitivity
68	YLR326W		Hypothetical ORF

5.4 Discussion

Microarray analysis has extensively been employed to identify targets of transcription factors or to study gene expression patterns under different environmental conditions. The great advantage of microarray analysis is that the expression patterns of virtually all genes of a genome can be investigated in one experiment. We applied this method to explore novel functions of the essential protein phosphatase Glc7p. Interestingly, we found that in *glc7-12* mutants transcription of a large number of RP and Ribi genes is downregulated at the restrictive temperature compared to the WT. Comparison to the WT reveals which genes are up- or downregulated and is important to exclude effects that are merely due to the heatshock treatment. The expression profiles of several other genes, such as Trs31, Nrd1 and Rec104 gave first indications that Glc7p could also be involved in snoRNA transcription termination (described in more detail in Chapter 3). Thus, this microarray analysis was instrumental in the discovery of new cellular functions of Glc7p.

Analysis of the microarray data

Because Glc7p is not a transcription factor that directly influences transcription of genes but a protein phosphatase, it is more intricate to determine what could have been the cause of an altered gene expression pattern in the *glc7-12* mutant. Protein phosphatase Glc7p has many different functions in the cell that not necessarily ever affect gene transcription. However, many pathways in a cell are regulated by phosphorylation that do ultimately impinge on the transcription of specific genes. Specific changes in the cell are expected to result in a common transcriptional response of genes belonging to a functional group. For example, genes that encode proteins which are required for the synthesis of a ribosome should all respond to an environmental change in the same way. It was therefore instrumental to cluster genes according to their expression pattern to identify specific cellular processes affected by Glc7p. We divided the 826 genes that were differentially expressed in *glc7-12* mutants into six clusters. Each cluster contains the genes that have an expression pattern that is more similar to the genes in their cluster than to genes in any other cluster. Three of the clusters contained genes that were upregulated and three contained genes that were downregulated in *glc7-12* mutants. We became most

interested in a subset of genes in cluster 1 and cluster 4, which are RP and Ribi genes. Both clusters comprise genes that are downregulated in *glc7-12* mutants. It has been reported that the transcription factor Sfp1p is involved in the regulation of both, the RP gene and the Ribi gene regulon (Jorgensen et al., 2004). In conclusion, Sfp1p could be a prime target for Glc7p in the regulation of RP and Ribi gene transcription. One of the reasons why RP and Ribi genes are not detected in the same cluster could be that regulation of RP gene transcription involves an additional number of transcription factors. These could integrate different signal inputs and evoke a slightly different transcriptional response. Thus, the expression pattern of RP genes could be distinct from the one of Ribi genes.

Microarray analysis in our hands was a great tool to gain a genome wide overlook of transcriptional responses under certain conditions. It is however recommended to cross-check microarray results by independent approaches such as Northern analysis.

6 Glc7p is involved in transcription regulation of ribosomal protein genes

6.1 Summary

The transcription of genes that encode for factors required for ribosome biogenesis is tightly controlled and responds to extracellular and intracellular changes. The target of rapamycin (TOR) and the Ras/PKA signaling pathways both regulate transcription of RP and Ribi genes in response to nutrient conditions. Here, we show that the essential protein phosphatase Glc7p is involved in regulation of RP and Ribi gene transcription. Firstly, transcription of RP and Ribi genes is downregulated in *glc7* mutant strains as observed by both, microarray and northern analysis. Secondly, Glc7p controls the nuclear localization of two downstream targets of PKA, Yak1p and Crf1p, which have been implicated in RP gene transcription regulation. In addition, we show that PKA is active in glc7 mutant strains and that Glc7p most likely acts downstream of PKA or in parallel to Yak1p. Furthermore, Glc7p regulates the phosphorylation level of a specific isoform of Bcy1p. This regulation however, is not required in the regulation of RP gene transcription. Therefore, the direct target of Glc7p in the transcription regulation of RP and Ribi genes remains unknown.

6.2 Introduction

Growing cells increase the number of ribosomes to adjust to the rising demand for proteins. Therefore, regulation of ribosome synthesis is crucial for the control of cell growth. Ribosomes consist of ribosomal proteins (RP) and ribosomal RNA (rRNA). There are 138 genes encoding 79 different RPs in *S. cerevisiae*. Transcription of RP genes is regulated coordinately in response to growth stimuli and environmental stress (Warner, 1999). RP gene transcription is stimulated under favorable growth conditions and approximately 50% of RNAP II transcription is devoted to RP genes (Warner, 1999). Under non-favorable conditions RP gene transcription is rapidly and coordinately repressed (Warner, 1999). A number of transcription factors have been identified that regulate the transcription of RP genes. One of them, Rap1p, constitutively binds to Rap1 binding sites present at most RP gene promoters (Cardenas et al., 1999; Warner, 1999). However, Rap1p also binds and activates many

non-RP promoters in a manner not coordinated with those of RP gene promoters. These include glycolytic genes (Morse, 2000). Rap1p also acts as the major duplex DNA binding protein of telomeres (Morse, 2000; Pina et al., 2003). Therefore, to achieve specificity, Rap1p works together with specific cofactors to activate transcription of RP genes. Such cofactors have recently been identified as Fhl1p, Ifh1p and Crf1p (Martin et al., 2004; Schawalder et al., 2004; Wade et al., 2004). A global chromatin localization analysis in yeast revealed the forkhead-like protein Fhl1p as a factor highly specific for RP gene promoters (Lee et al., 2002). Ifh1 is recruited to RP gene promoters by binding to Fhl1p (Martin et al., 2004; Schawalder et al., 2004; Wade et al., 2004). Together these factors activate the transcription of RP genes. Crf1p was described as a co-repressor of RP gene transcription that competes with Ifh1p for binding to Fhl1p under non-favorable nutrient conditions (Martin et al., 2004). Both, the association of Fhl1p and Ifh1p are dependent on Rap1p (Wade et al., 2004). Binding of Crf1p to Fhl1p represses transcription of RP genes (Martin et al., 2004). However, the potential of Crf1p to repress transcription of RP genes might be restricted to certain strain-backgrounds (Zhao et al., 2006) and thus may not be part of an essential ”core” repression mechanism. The current model is that Rap1p and Fhl1p are constitutively bound at RP gene promoters, however activation of RP gene transcription is achieved only by further binding of Ifh1p to Fhl1p. This suggests that Ifh1p plays a central role in the regulation of RP gene transcription (Martin et al., 2004; Rudra et al., 2005; Schawalder et al., 2004; Wade et al., 2004).

A large number of genes called Ribi genes, show nearly identical transcriptional responses to environmental perturbations as do RP genes (Jorgensen et al., 2002; Wade et al., 2001). Most of the Ribi genes encode proteins that are accessory in ribosome synthesis and modify rRNA and RPs in the nucleolus (Fatica and Tollervey, 2002; Jorgensen et al., 2002; Wade et al., 2001). The promoters of Ribi genes are strongly enriched for two motifs, which are absent in RP gene promoters, termed RRPE and PAC. Therefore, Ribi gene promoters appear to constitute a distinct regulon (Jorgensen et al., 2002; Wade et al., 2001). The transcription factor Sfp1p, previously shown to play a role in the control of cell size (Jorgensen et al., 2002), is an activator of the RP and Ribi gene regulons (Jorgensen et al., 2004; Marion et al., 2004). The nuclear concentration of Sfp1p responds rapidly to changes in nutrients and environmental conditions. Under optimal growth conditions, Sfp1p is localized in the nucleus (Jorgensen et al., 2004).

Two central nutrient-sensing signaling pathways can activate the transcription of rRNA, RP and Ribi genes: the target of rapamycin (TOR) and the Ras/PKA signaling pathway (Cardenas et al., 1999; Klein and Struhl, 1994; Neuman-Silberberg et al., 1995; Powers and Walter, 1999; Wang et al., 2004). These two signaling pathways control the localization of Rap1p, Fhl1p, Ifhl1p and Crf1p to RP and Ribi gene promoters (Martin et al., 2004).

Core components of the Ras/PKA signaling pathway are the RAS-GTPase(s), the adenylate cyclase and the cAMP-dependent protein kinase A (PKA). Ras proteins belong to a superfamily of GTP-bound proteins (Broach and Deschenes, 1990). In *S. cerevisiae*, the Ras proteins are encoded by the RAS1 and RAS2 genes (Broach and Deschenes, 1990). Neither Ras1p nor Ras2p is essential for cell growth, however deletion of both genes is lethal (Broach and Deschenes, 1990). Ras proteins are localized at the plasma membrane were they receive information about the nutritional/environmental extracellular conditions (Broach and Deschenes, 1990). Ras proteins are activated by Cdc25p, which promotes the exchange of GDP to GTP within the Ras-GTPase (Lai et al., 1993; Robinson et al., 1987). Ira1p and Ira2p function as GAPs that stimulate the RAS GTPase leading to inactivation of Ras (Tanaka et al., 1990a; Tanaka et al., 1990b). When activated, Ras activates the adenylate cyclase, encoded by the CDC35 gene (Toda et al., 1985). Adenylate cyclase converts ATP to the second messenger cAMP (Toda et al., 1985). cAMP in turn activates PKA. PKA is a heterotetramer consisting of two inhibitory subunits encoded by BCY1 and two catalytic subunits, redundantly encoded by TPK1, TPK2 and TPK3 (Toda et al., 1987a; Toda et al., 1987b). Binding of cAMP to Bcy1 leads to the dissociation of the catalytic subunits from Bcy1p. The catalytic subunits of PKA can then phosphorylate their different cellular targets. One of these targets is the kinase Yak1p. PKA also negatively regulates the transcription of many stress-responsive genes by phosphorylating and thereby inactivating the transcription factor Msn2 (Gorner et al., 1998). Mutations that render Bcy1 inactive still yield conditionally active PKA. Cells that carry such a mutation fail to accumulate carbohydrates and do not arrest in the G1 phase upon nutrient limitation. In addition, regulatory enzymes of glycolysis and of gluconeogenesis are stimulated or inhibited, respectively, by high PKA activity. Therefore, it was proposed that the Ras/PKA signaling pathway communicates nutritional conditions to the cell (Thevelein and de Winde, 1999). Inactivation of Ras or PKA or depletion of cAMP leads to the arrest of yeast cells in

G1 of the cell cycle. Interestingly, nutrient-starved cells arrest at the same point in G1. Subsequently cells with inactivated Ras/PKA signaling enter the stationary phase (G0) and display physiological changes normally associated with nutrient limitation. These changes include an altered transcription pattern and changes in expression of RP genes, the accumulation of carbohydrates, and also increased resistance towards heat and oxidative stress (Thevelein and de Winde, 1999; Werner-Washburne et al., 1996).

The conserved phosphatidylinositol kinase TOR controls many aspects of cell growth. In yeast two Tor proteins exist, Tor1p and Tor2p, which are highly homologous (Helliwell et al., 1994). They stimulate cap-dependent translation initiation, membrane traffic, protein degradation, ribosome biogenesis and transcription in response to nutrients. TOR kinases are Ser/Thr protein kinases. Tor1p and Tor2p are inhibited by the immunosuppressant and antibiotic rapamycin, which binds to a domain adjacent to the catalytic kinase domain (Chen et al., 1995). Rapamycin treatment or TOR depletion results in several physiological changes characteristic of starved (G0) cells, including accumulation of storage carbohydrates and an altered transcription pattern. This includes downregulation of RP gene transcription (Barbet et al., 1996). The striking similarities between starved and rapamycin-treated cells suggest that the TOR signaling pathway controls cell growth in response to nutrients (Barbet et al., 1996; Thomas and Hall, 1997). How the nutrient signal is transduced to TOR however is still unknown. Under favorable nutrient conditions TOR globally represses transcription of starvation-specific genes by preventing nuclear localization of several nutrient-responsive transcription factors (Crespo and Hall, 2002). How TOR impinges on RP gene transcription has only recently become clearer. As described above, Fhl1p, Ifh1p and Crf1p were identified as transcription factors that among others control the transcription of RP genes. Several reports indicated that TOR controls the association of Ifh1p and Crf1p with RP gene promoters leading to activation or repression of RP genes, respectively (Martin et al., 2004; Schawalder et al., 2004; Wade et al., 2004; Figure 6.1). Inhibition of the TOR pathway by rapamycin causes a rapid dissociation of Ifh1p, but not Fhl1p or Rap1p from RP gene promoters (Martin et al., 2004; Schawalder et al., 2004; Wade et al., 2004). In addition, Crf1p associates with RP promoters under these conditions (Martin et al., 2004). Moreover, in response to inhibition of TOR, also Sfp1p is released from RP gene promoters and

leaves the nucleus leading to downregulation of RP gene transcription (Marion et al., 2004).

Connections between the TOR and the Ras/PKA pathways have been described (Schmelzle et al., 2004). For instance, the rapamycin-induced downregulation of RP gene transcription can be counteracted by activation of the Ras/PKA pathway (Schmelzle et al., 2004). In addition, TOR controls the subcellular localization of PKA and the PKA target Yak1p and a yak deletion confers rapamycin resistance (Schmelzle et al., 2004). Moreover, Yak1p was shown to phosphorylate Crf1p in cells treated with rapamycin leading to nuclear localization of Crf1p and RP gene repression (Martin et al., 2004). These findings led to the proposition that the Ras/PKA pathway might be a novel TOR effector branch (Schmelzle et al., 2004).

Here, we report, that the essential protein phosphatase Glc7p is involved in the regulation of RP and Ribi gene transcription. Glc7p controls the nuclear localization of Crf1p and Yak1p. In addition, in studies using the hyperactive Ras2^{val19} allele and localization experiments with the PKA subunits Bcy1p and Tpk1p suggested that Glc7p acts downstream of PKA or in parallel to Yak1p. Interestingly, we found that Glc7p regulates the phosphorylation state of one of the Bcy1 phospho-isoforms. However regulation of this specific phospho-isoform does not affect RP gene transcription.

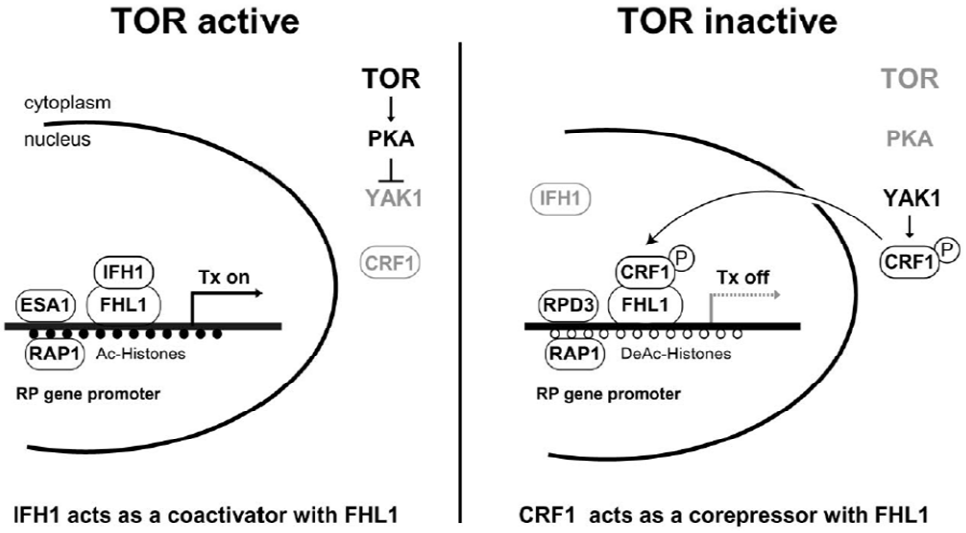


Figure 6.1. TOR regulates ribosomal protein gene expression via PKA and the granscription factor Fhl1 (Martin et al., 2004).

6.3 Results

6.3.1 Mutations in Glc7p lead to the downregulation of RP and Ribi gene transcription

To find out whether the expression of specific subsets of genes are affected in a *glc7* mutant, we performed expression profiling of poly(A)$^+$ RNA derived from *glc7-12*, a temperature sensitive *glc7* allele, and the WT (chapter 5; Materials and Methods).
We analyzed our data using a stringent p-value of < 0.001. This means that the probability that the results observed could have occurred by chance is less than 1‰. We found 826 genes to be differentially expressed of which 417 were upregulated and 409 were downregulated. We noticed that the latter group contained a large amount of ribosomal protein (RP) and Ribi genes, namely 36 RP and 57 Ribi genes (Table 6.1). Transcription of RP mRNA represents ~50% of the total RNAP II transcription initiation events, which consume much of the cells energy and therefore must be tightly controlled. *S. cerevisiae* has 138 genes encoding for 79 different ribosomal proteins (Zhao et al., 2006) and 236 Ribi genes (Jorgensen et al., 2004). Analysis of the raw data of our microarray revealed that all RP genes were downregulated in *glc7-12* under non-permissive conditions, indicating that a p-value of <0.001 is highly stringent. The mRNA of RP and Ribi genes is very unstable (Grigull et al., 2004). As a result, a decrease in the mRNA level of these genes is indicative of a failure to initiate transcription. We therefore assume that Glc7p is involved in the regulation of RP and Ribi gene transcription.

Table 6.1

RP genes downregulated in *glc7-12*

YORF	Gene name
RPL	
YPL220W	RPL1A
YFR031C-A	RPL2A
YOR063W	RPL3
YBR031W	RPL4A
YPL131W	RPL5
YLR448W	RPL6B
YGL076C	RPL7A
YPL198W	RPL7B
YHL033C	RPL8A
YLL045C	RPL8B
YLR075W	RPL10
YDR418W	RPL12B
YKL180W	RPL17A
YJL177W	RPL17B
YNL301C	RPL18B
YPL079W	RPL21B
YLR061W	RPL22A
YLR325C	RPL38
RPS	
YGR214W	RPS0A
YLR048W	RPS0B
YGL148W	RPS2
YNL178W	RPS3
YJR123W	RPS5
YBR189W	RPS9B
YBR048W	RPS11B
YDR064W	RPS13
YOL040C	RPS15
YNL302C	RPS19B
YJL190C	RPS22A
YLR367W	RPS22B
YGL189C	RPS26A
RPP	
YDL081C	RPP1A
YDL130W	RPP1B
YOL039W	RPP2A
YDR382W	RPP2B
Mitochondrial RP genes	
YDR237W	MRPL7
YBR146W	MRPS9

Table 6.1 continued

Ribi genes that are downregulated in *glc7-12*

YORF	Gene name
Nucleotide metabolism	
YML056C	IMD4
YHL011C	PRS3
YMR217W	GUA1
YKL181W	PRS1
YBL039C	URA7
YML106W	URA5
YKL216	URA1
YNL141W	AAH1
YDR399W	HPT1
YER056C	FCY2
YML022W	APT1
YHR128W	FUR1
YNR012W	URK1
YPR074C	TKL1
Ribosome biogenesis	
YHR169W	DBP8
YNL112W	DBP2
YBR267W	-
YHR089C	GAR1
YPL093W	NOG1
YGL120C	PRP43
YDR101C	ARX1
YMR049C	ERB1
YPR144C	NOC4
YOR145C	PNO1
YMR290C	HAS1
YCL059C	KRR1
YMR093W	UTP15
YJL109C	UTP10
YOL077C	BRX1
YLR002C	NOC3
YMR116C	ASC1
YOR272W	YTM1
YPL211W	NIP7
YGL111W	NSA1
YPL266W	DIM1
YLR129W	DIP2
YGR159C	NSR1
YNR038W	DBP6
YMR239C	RNT1

YORF	Gene name
RNA Polymerase I and III	
YOR224C	RPB8
YNR003C	RPC34
YOR341W	RPA190
YNL248C	RPA49
YPR010C	RPA135
tRNA synthetases or metabolism	
YGR264C	MES1
YOL097C	WRS1
YDR165W	TRM82
YGR169C	PUS6
YPR033C	HTS1
Translation initiation and elongation factors	
YNL062C	GCD10
YER025W	GCD11
YFR009W	GCN20
YKR026C	GCN3
YKL081W	TEF4
YJL138C	TIF2
YMR146C	TIF34
YDR429C	TIF35
YJR007W	SUI2
YLR249W	YEF3

6.3.2 Northern analysis of a subset of RP and of one Ribi gene confirms the downregulation of these genes in several glc7 mutant alleles

To confirm that RP and Ribi gene transcription was downregulated in *glc7-12* mutants, we carried out Northern analysis with probes against several RP mRNAs (Rps11, Rps13, Rps16, Rpl 28) and the mRNA of Asc1, a Ribi gene (Fig. 6.2). Rps11, Rps13 and Asc1p were among the RP and Ribi genes identified in the microarray

(p<0,001), whereas Rps16 and Rpl28 were not. We extracted total RNA from *glc7-12* at the permissive and the non-permissive temperature. In addition, we collected total RNA from several other temperature sensitive *glc7* mutant strains to find out whether the transcriptional downregulation of RP genes would be allele-specific. Whereas *glc7-10* showed a modest decrease in mRNA levels, *glc7-12*, *glc7-13* and *glc7-5* showed a strong or complete loss of mRNAs of all the mRNAs tested. The fact that Rps 16 and Rpl 28 mRNA levels are decreased as clearly as levels for Rps11 and Rps13 reinforces the notion that downregulation applies to all RP genes. Therefore, this Northern analysis shows that we applied very stringent conditions in our microarray that do not reveal the complete set of genes that are differentially expressed in *glc7-12*. This also indicates that the microarray data are highly significant. Furthermore, we conclude that transcription of RP genes is repressed in several *glc7* mutant strains.

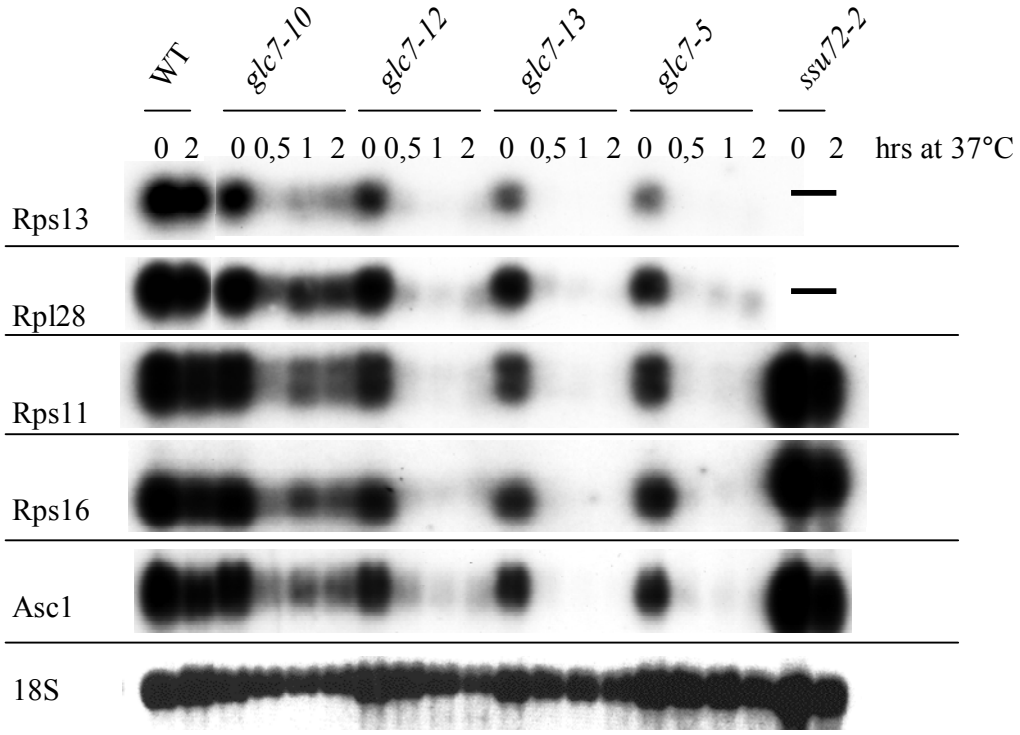


Figure 6.2. Expression of ribosomal protein genes is downregulated in glc7 mutants.
Northern analysis of 20µg total RNA extracted from WT, *glc7-10*, *glc7-12*, *glc7-13, glc7-5* and *ssu72-2* cells following growth at RT or after shift to 37º C for the indicated times. Filters were probed with random-primed labeled oligonucleotides directed against the RNAs indicated on the left. 18S served as loading control.

6.3.3 The CTD of RNAP II is most likely not involved in the transcription regulation of RP and Ribi genes

Transcription of genes is regulated at many levels. We wanted to exclude the possibility that the downregulation of RP and Ribi genes results from a general transcription failure due to misregulation of the transcription apparatus. The CTD of RNAP II constitutes a "landing platform" for factors that associate with it and co-transcriptionally process the nascent transcript (Proudfoot, 2000a). However, the function of these processing factors is not limited to the production of mature mRNAs, but they also influence the transcriptional activity of RNAP II (Zorio and Bentley, 2004). By providing a platform for transcription processing factors the CTD is an important link in the communication between RNAP II transcription and maturation of the transcripts.

The highly dynamic phosphorylation pattern of the CTD ensures that specific processing factors bind the CTD at specific phases in the transcription cycle (see chapter 1.6). The phosphorylation pattern of the CTD is achieved by a set of kinases and phosphatases. Among the know phosphatases is Ssu72, a component of CPF. Ssu72 dephosphorylates serine-5 of the RNAP II CTD to release RNAP II from a transcriptional pause site and to promote productive transcription (Krishnamurthy et al., 2004).

We asked whether interfering with the CTD phosphorylation pattern would disturb RP gene transcription. Therefore, we were interested if the mRNA levels of RP and Ribi genes would decrease in the *ssu72-2* mutant. However, we detected normal levels of the respective mRNAs (Rps11, Rps16 and Asc1) in *ssu72-2* after shift to the non-permissive temperature (Fig. 6.2). We did not test Rps13 and Rpl28 in this mutant. This indicates that a disturbed CTD phosphorylation pattern is not sufficient to cease transcription of RP genes. Moreover, we and others (He and Moore, 2005) could not find any proof that Glc7p dephosphorylates the CTD *in vivo* (see Fig.3.8 chapter 3.3.10). Altogether, this could suggest that the CTD is not the target of regulation in the control of RP gene transcription.

6.3.4 Crf1 accumulates in the nucleus of glc7-12 and glc7-13 mutants under non-permissive conditions

Recently, it has been shown that a set of transcription factors consisting of Fhl1p, Ifh1p and Crf1p, specifically regulate the transcription of RP genes (Martin et al., 2004; Schawalder et al., 2004; Wade et al., 2004). Fhl1 acts in concert with the co-activator Ifh1p or the co-repressor Crf1p to activate or repress the transcription of RP genes. The binding of Crf1p to Fhl1p represses transcription of RP genes.

Rapamycin treatment of cells, which mimics nutrient depletion, leads to a downregulation of RP gene transcription. One consequence of rapamycin treatment is the rapid import of Crf1p from the cytoplasm to the nucleus. While phosphorylation of Crf1p is crucial for nuclear localization it is probably not needed for the interaction with Fhl1p (Martin et al., 2004).

We asked whether mislocalization of Crf1p to the nucleus might be one of the reasons for the downregulation of RP genes in our *glc7* mutants. To analyze the sub cellular localization of Crf1p, we carried out indirect immunofluorescence experiments with cells expressing HA-tagged Crf1p. In contrast to WT, we see a fast and strong accumulation of HA-Crf1p in the nucleus in *glc7-12* and *glc7-13* mutant cells after shift to the non-permissive temperature (Fig. 6.3). In *glc7-13* we readily detect HA-Crf1p in the nucleus after a 15 min shift to 33º C. Also *glc7-12* accumulates HA-Crf1p in the nucleus after a 30 min shift to 37ºC. We did not test shorter time points for *glc7-12*.

These experiments show that inactivating the phosphatase activity of Glc7p by heat has the same effect on the localization of Crf1p as rapamycin addition. We can exclude that this effect is due to the heat shock, because WT cells do not accumulate Crf1p in the nucleus.

Taken together these experiments suggest that the high concentration of Crf1p in the nucleus could contribute to the repression of RP gene transcription which we observe in these mutants. However, while it was clearly demonstrated that Crf1p acts as a co-repressor of RP gene transcription when using strains with an TB50 background (Martin et al., 2004), the W303 strain background used in this study does not depend on Crf1p. It has recently been reported that Crf1p does not associate with the promoter region of an RP gene and that deletion of Crf1p does not abolish repression of RP gene transcription by rapamycin in this strain background (Zhao et al., 2006).

Since we work in a W303 background, a function of Crf1p in RP gene repression may not be a general phenomenon. Nevertheless, this experiment indicates that nuclear accumulation of Crf1p can be controlled by Glc7p.

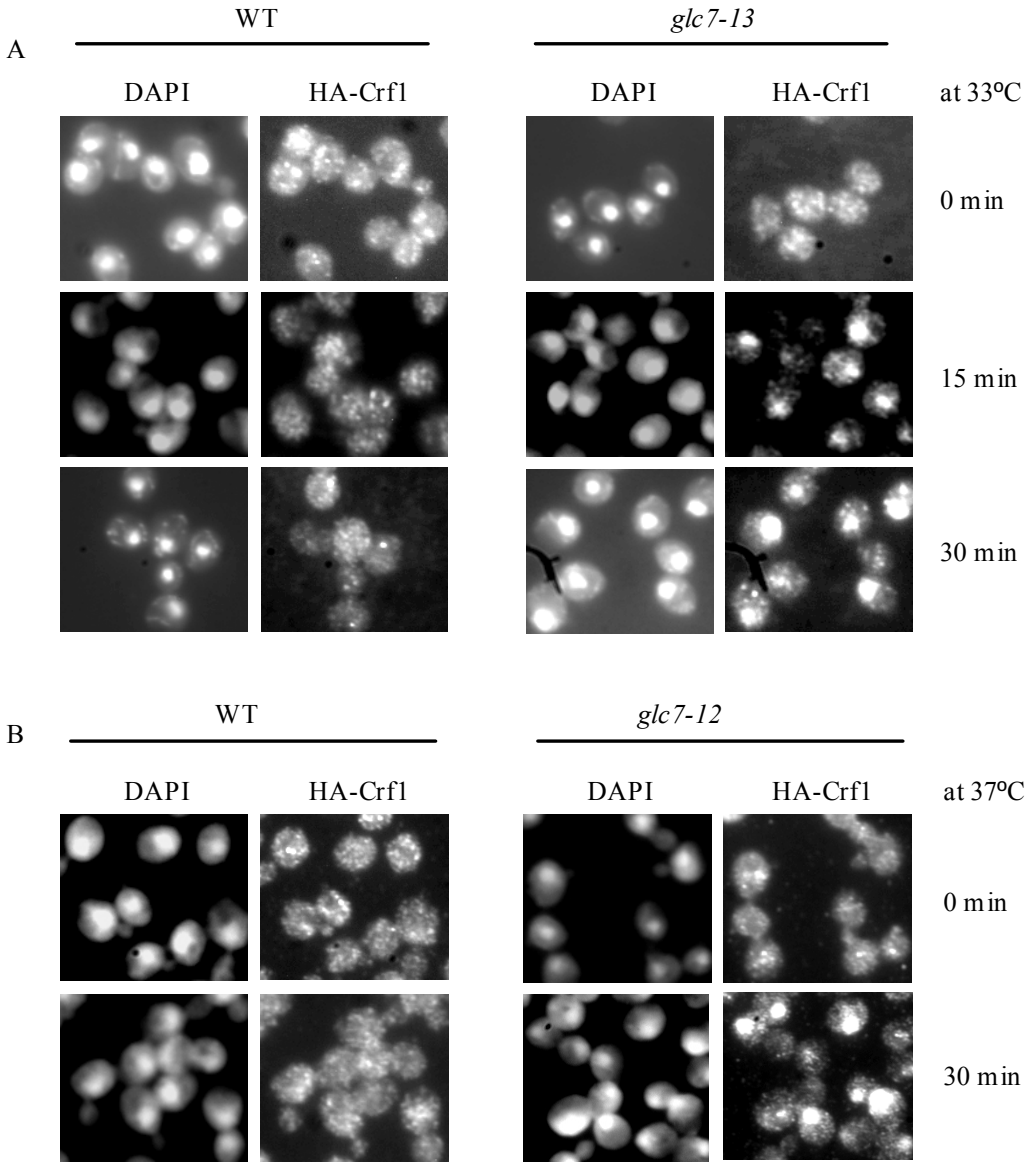


Figure 6.3. Glc7 controls CRF1 nuclear localization
A) Cells were grown in SD medium to OD_{600} 0.8 and shifted to 33ºC for the indicated times. CRF1 localization was examined by indirect immunofluorescence on whole fixed cells. In all cases, plasmid-borne CRF1 was tagged with the HA epitope and a monoclonal anti-HA antibody was used as a primary antibody for indirect immunofluorescence. Strains were as follows: WT and *glc7-13* transformed with pDM40. B) Experimental procedures as in A) except cells were shifted to 37ºC for the indicated times. Strains were as follows: WT and *glc7-12* transformed with pDM40.

6.3.5 Nuclear accumulation of Yak1p is regulated by Glc7p

Under non-favorable nutrient conditions the kinase Yak1p phosphorylates Crf1p. Subsequently, Crf1p translocates to the nucleus and competes with Ifh1p for binding to Fhl1p (Martin et al., 2004). Yak1p kinase activity is regulated negatively, probably directly, by PKA (Garrett and Broach, 1989; Griffioen et al., 2001; Moriya et al., 2001; Zappacosta et al., 2002). Therefore, the activity of Yak1p is greatly enhanced when cells are treated with rapamycin or when starved for glucose. In addition, these conditions lead to nuclear accumulation of Yak1p (Moriya et al., 2001; Schmelzle et al., 2004). In an attempt to test whether nuclear localization of Crf1p was caused by an enhanced activity of its kinase Yak1p, we looked at the sub cellular localization of Yak1p, as it accumulates in the nucleus when it is activated. To this end, we performed indirect immunofluoresence microscopy on WT and *glc7-13* cells expressing plasmid-borne HA-tagged Yak1p. We found that Yak1p accumulates in the nucleus of *glc7-13* cells but not in WT cells after a shift to the non-permissive temperature of 33º C for 30 min (Fig. 6.4, WT not shown).

This indicates that Yak1p is active in our mutant strain at 33ºC and is therefore able to promote nuclear localization of Crf1p likely by phosphorylating Crf1p.

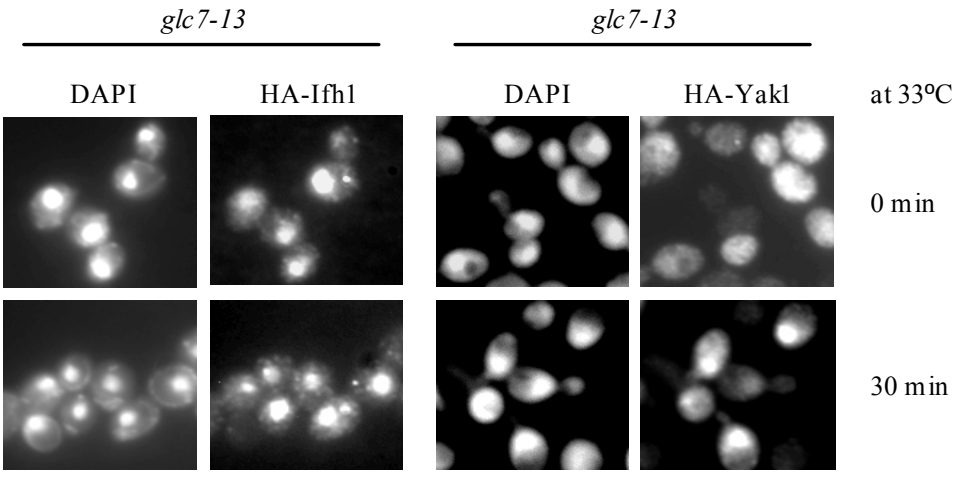


Figure 6.4. Yak1 accumulates in the nucleus in *glc7-13* mutant cells
Cells were grown in SD medium to OD_{600} 0.8 and shifted to 33ºC for the indicated times. YAK1 and IFH1 localization was examined by indirect immunofluorescence on whole fixed cells. Plasmid-borne YAK1 and IFH1 was tagged with the HA epitope and a monoclonal anti-HA antibody was used as a primary antibody for indirect immunofluorescence. Strains were as follows: *glc7-13* transformed with pDM55 (IFH1; left panel) or pDM57 (YAK1; right panel). IFH1 served as a nuclear marker.

6.3.6 *Constitutively active Ras signaling suppresses RP gene transcription failure in glc7-12 but not in glc7-13*

In *S. cerevisiae*, the RAS/cAMP and the TOR signaling pathway play important roles in communicating nutrient conditions to the cell. It has been shown that the two pathways share the regulation of RP gene transcription, because the hyperactive $Ras2^{val19}$ allele of Ras2, is able to suppress a TOR deficiency and restores RP gene transcription after rapamycin treatment (Schmelzle et al., 2004). Consequently, under favorable nutrient conditions both pathways support the transcription of RP genes. We were interested if $Ras2^{val19}$ could suppress the downregulation of RP gene transcription in *glc7* mutants. Therefore, we transformed *glc7-12* and *glc7-13* with a plasmid expressing $Ras2^{val19}$ or an empty vector control and extracted total RNA from cells before and after shift to the non-permissive temperature. RP gene transcription is transiently downregulated in response to heat shock but thereafter increases constantly until reaching normal levels (see results for WT with empty vector control; Fig. 6.5 a, D. Martin unpublished). Hyperactive Ras signaling however suppresses this transient downregulation in the WT and also in *glc7-12* mutant cells. Interestingly, hyperactive Ras signaling does not result in transcription of RP genes in *glc7-13* at the non-permissive temperature, indicating that Glc7p acts downstream or in parallel to Ras/TOR. However, we observed the complete opposite for *glc7-12,* which suggests that Glc7p acts upstream of Ras and/or TOR. We suspect that there could be more than one function for Glc7p in these signaling cascades. In addition, Glc7p could act independent of this signaling. Taken together, we conclude that Glc7p acts downstream or in parallel to Ras/TOR and in addition functions independent of this signaling pathway to regulate the transcription of RP genes.

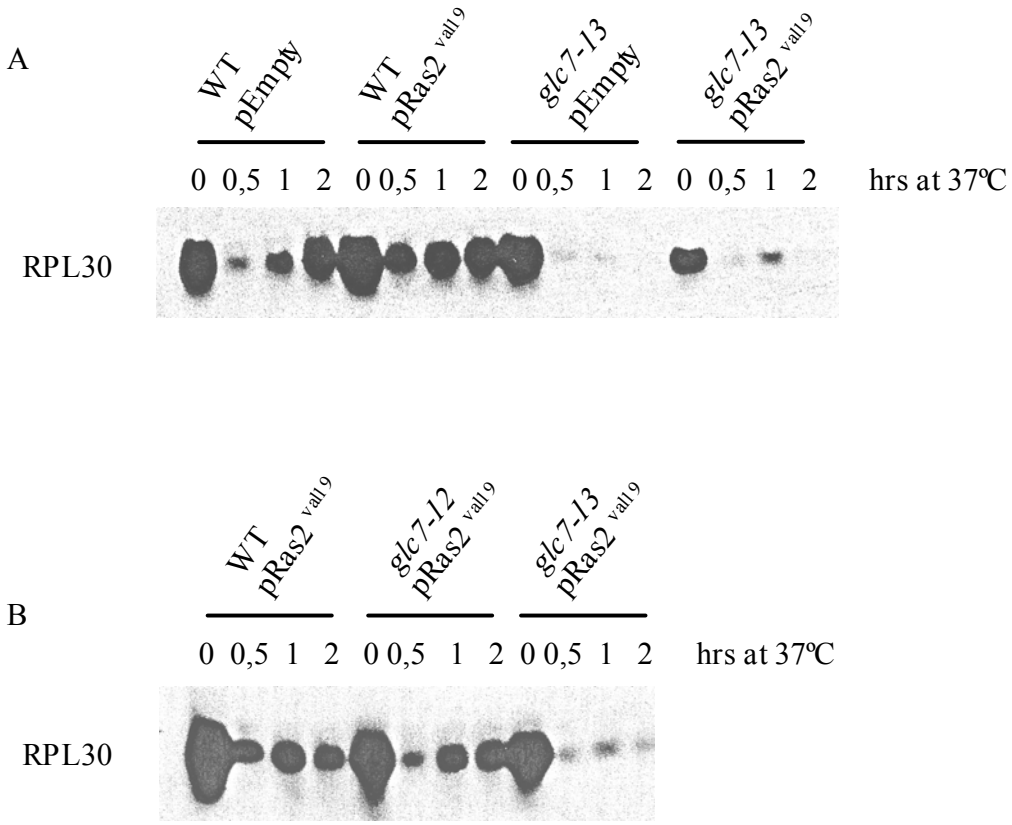


Figure 6.5. Allele-specific effect of Ras2^{val19} on expression of RPL30 in *glc7* mutant cells
WT, *glc7-12* and *glc7-13* cells were transformed with pTS 118 (Ras2^{val19}) or with an empty vector control. Northern analysis of 20µg total RNA extracted from WT, *glc7-12*, *glc7-13* cells was performed following growth at RT or after shift to 37º C for the indicated times. Filters were probed with random-primed labeled oligonucleotides directed against RPL30. A) Ras2^{val19} does not suppress downregulation of RPL30 in *glc7-13* mutant cells. B) Ras2^{val19} does suppress downregulation of RPL30 in *glc7-12* mutant cells.

6.3.7 PKA is active in glc7-12 and glc7-13 suggesting that Glc7p acts downstream of PKA or in parallel to Yak1

From the experiments with the hyperactive Ras ^{val19} allele we could not firmly conclude that Glc7p is part of the Ras/PKA and the TOR signaling pathways. We therefore decided to investigate PKA. The kinase PKA was shown to be important for signal transmission to the nucleus in the TOR as well as in the cAMP/Ras signaling pathway. Importantly, PKA negatively regulates the kinase Yak1p (Martin et al., 2004), which is active in our *glc7-13* mutant cells. Since Yak1p is active in *glc7-13* under non-permissive conditions this would imply that PKA itself is inactive.

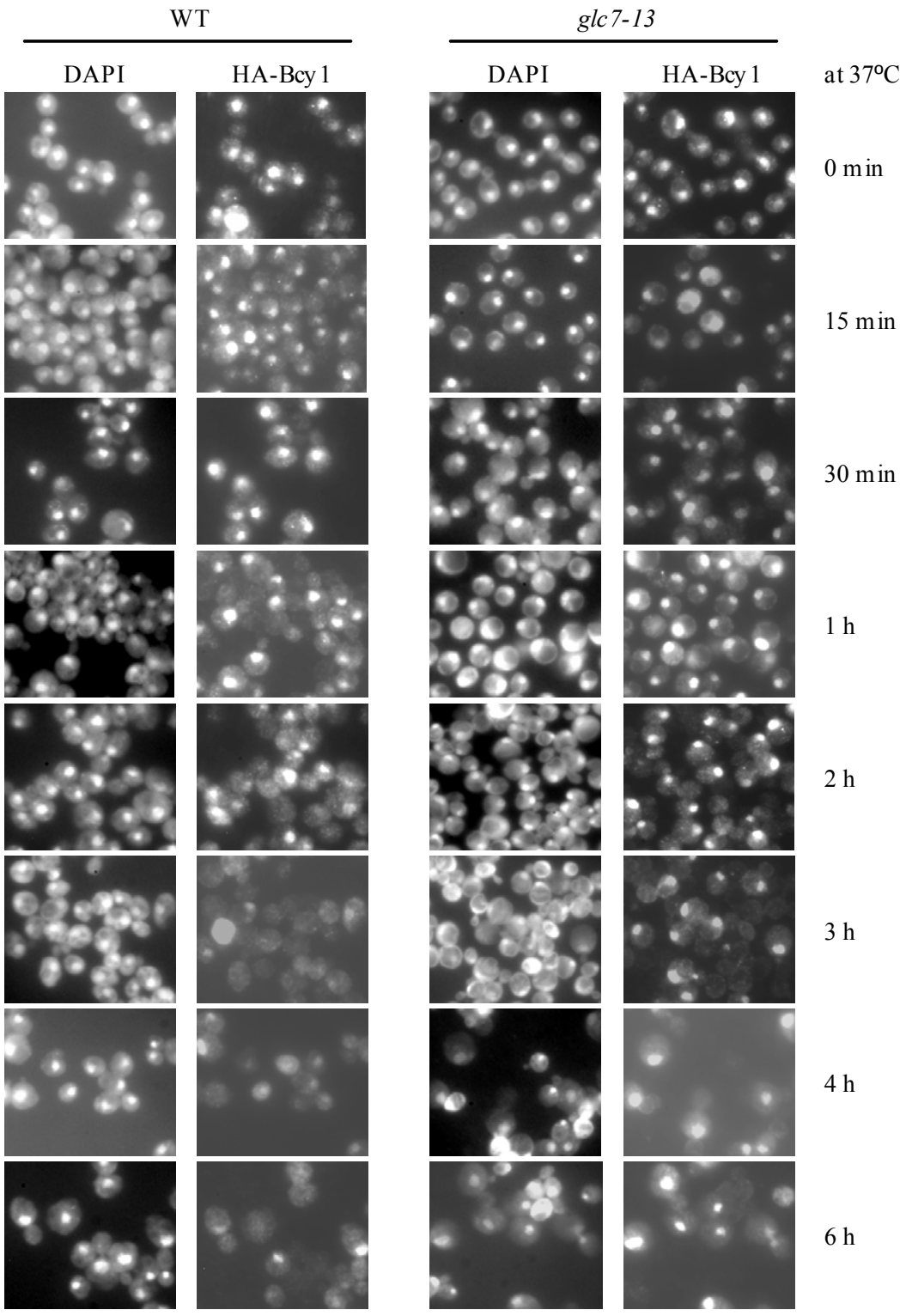
PKA consists of a catalytic subunit, redundantly encoded by the genes Tpk1p, Tpk2p and Tpk3p (Toda et al., 1987b) and the regulatory subunit Bcy1p. Dissociation of Bcy1 from Tpk1p leads to activation of the catalytic subunit, which takes place when cAMP binds to Bcy1p. cAMP production is triggered by glucose. At least two different mechanisms determine the sub cellular localization of PKA: first, cAMP controls the localization of Tpk1p. When cAMP is abundant Tpk1 is distributed equally all over the cell. If cAMP levels drop, Tpk1 accumulates in the nucleus. Second, the carbon source determines the localization of Bcy1p. Whereas Bcy1p is predominantly nuclear in cells grown on glucose, it is also found in the cytoplasm in cells grown on ethanol (Griffioen et al., 2000). In addition, rapamycin induces the nuclear localization of both Tpk1p and Bcy1p. To understand the function of Glc7p in the regulation of RP gene transcription, it was important to see whether PKA was active in our *glc7* mutants and if localization of PKA would give any hint for the involvement of Glc7p in either the TOR or cAMP/Ras signaling pathways or both.

We first looked at the cellular distribution of HA-tagged Bcy1 by indirect immunofluoresence microscopy and followed it over six hours at the non-permissive temperature in *glc7-13* cells. In agreement with published reports (Griffioen et al., 2003), we find that after shift to the non-permissive temperature Bcy1p is redistributed throughout the cell in WT cells (Fig. 6.6 a). In contrast, Bcy1 remained in the nucleus in *glc7-13* cells. This suggests that inactivating *glc7-13* has the same effect as rapamycin treatment on the intracellular localization of Bcy1p. This also suggests that Glc7p is part of one of the two signaling pathways.

Next, we compared localization of HA-tagged Tpk1p in WT and *glc7-13* cells after shift to the non-permissive temperature by indirect immunofluoresence microscopy. At RT Tpk1p is distributed equally within the cell in WT as well as in *glc7-13* cells (Fig. 6.6 b). Growing the cells at the non-permissive temperature leads to nuclear accumulation of Tpk1p in some cells in both the WT and in *glc7-13* mutant cells. This resembles the situation after rapamycin treatment. However, after one hour at the non-permissive temperature, Tpk1p did not show a predominantly nuclear localization anymore in *glc7-13* cells, but was detected all over the cell. In WT, many cells continued to show nuclear localization of Tpk1p. Therefore, in a time window of around half an hour at the restrictive temperature, WT and *glc7-13* mutant cells behave similar concerning the localization of Tpk1p. Within this time scale of half an hour however, we observe severe defects in RP gene transcription.

In conclusion, it is unlikely that Tpk1p is the target of Glc7p in the regulation of RP gene transcription. It is more likely that Glc7p regulates the activity of Tpk1p as a component of the glucose starvation signaling cascade because glucose starvation leads to nuclear export of Tpk1p. We conclude that Glc7p does not regulate Tpk1p activity via signal input of the TOR or cAMP/Ras signaling cascades but rather through signal input of the glucose starvation signaling cascade.

A


WT
glc7-13
DAPI
HA-Bcy1
DAPI
HA-Bcy1
at 37ºC
0 min
15 min
30 min
1 h
2 h
3 h
4 h
6 h

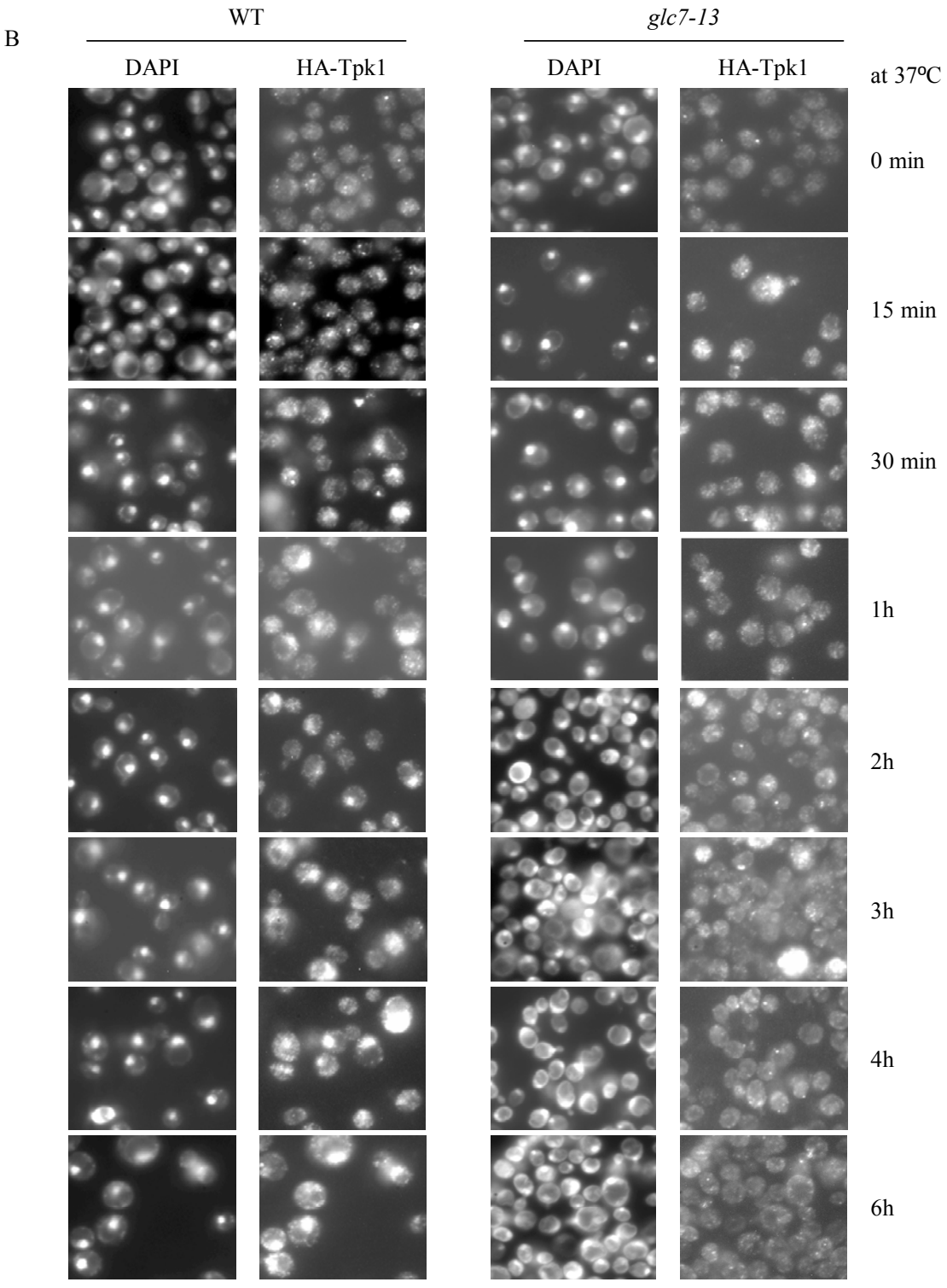


Figure 6.6. Glc7 controls nuclear localization of BCY1 and TPK1
A) Cells were grown in SD medium to OD_{600} 0.8 and shifted to 37ºC for the indicated times. BCY1 localization was examined by indirect immunofluorescence on whole fixed cells. Plasmid-borne BCY1 was tagged with the HA epitope and a monoclonal anti-HA antibody was used as a primary antibody for indirect immunofluorescence. Strains were as follows: WT and *glc7-13* transformed with pTS137. B) Experimental procedures as in A). Localization of TPK1 was examined in cells shifted to 37ºC for the indicated times. Strains were as follows: WT and *glc7-13* transformed with pTS134.

6.3.8 Glc7p controls phosphorylation of a distinct phosphoisoform of Bcy1p

The cytoplasmic localization of Bcy1 induced by heat stress requires phosphorylation of serines located in its localization domain (Griffioen et al., 2003). Interestingly these modifications do not affect PKA activity as such. In previous studies, Yak1p kinase was found to be required for cytoplasmic localization of Bcy1p in glucose deprived cells (Griffioen et al., 2001). However, it was suggested that this kinase is not important for temperature-dependent Bcy1p localization (Griffioen et al., 2003). In conclusion there seem to be two functionally independent phosphorylation events at the localization domain of Bcy1p that prompt Bcy1 to localize predominantly in the cytoplasm. To test whether mislocalization of Bcy1 in *glc7-13* cells to the nucleus at the non-permissive temperature was due to an alteration in the phosphorylation status of Bcy1p, we performed Western analysis of purified HA-tagged Bcy1p from WT, *glc7-12* and *glc7-13* cells, grown at the permissive and non-permissive temperature. An antibody directed against the HA-tag revealed the presence of slower migrating forms of Bcy1 within 15 min in the WT and in the mutant cells (Fig. 6.7).

Therefore, we conclude that phosphorylation of Bcy1 in response to heat stress is not significantly altered in *glc7-12* and *glc7-13* and thus cannot be the reason for the mislocalization of Bcy1 in *glc7-13* cells at 37ºC.

In addition to the phosphoisoform of Bcy1 induced by heat stress, there are at least two additional phosphoisoforms of Bcy1p visible in *glc7-12* and *glc7-13* after growth for 1 hour at 37ºC or more. These additional isoforms are only very weak in WT. This suggests that Glc7p regulates the formation of a distinct set of phosphoisoforms of Bcy1 and that this phosphorylation does not occur in response to heat stress. In addition, the kinase Yak1 is not responsible for this phosphorylation, since phosphorylation by Yak1p leads to cytoplasmic localization of Bcy1p. Finally, we suggest that the additional phosphorylation of Bcy1 observed in the *glc7* mutants, reverts the effect of the heat shock phosphorylation concerning localization of Bcy1p.

Western analysis of heat-stressed cells producing an HA-tagged version of Bcy1p with an antibody directed against the phosphorylated target region of PKA showed an increase in the signal in *glc7-12* and *glc7-13* compared to the WT. However, we consider these results to be preliminary because WT and mutant extracts were not assayed on the same gel (data not shown). Although we see a strong difference

between the WT and the mutants this could in part be due to variations between experiments. Nevertheless, these experiments suggest that Tpk1p is phosphorylating Bcy1p, but that the phosphate is not removed anymore in the mutants under conditions when Glc7p is inactive, meaning that Glc7p could be the responsible phosphatase. This also suggests that Tpk1 is active and responsible for the nuclear localization of Bcy1p to the nucleus.

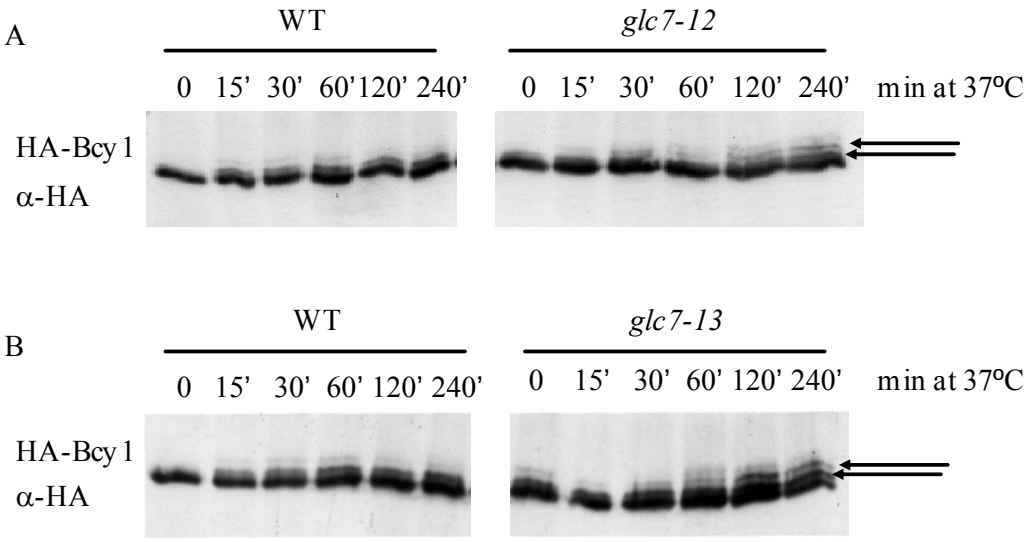


Figure 6.7. Specific phospho-isoforms of Bcy1 accumulate in *glc7* mutant cells
HA-tagged BCY1 was immunoprecipitated from WT, *glc7-12* (A) and *glc7-13* (B) cell extracts before and after shift to 37ºC for the indicated times. Purified HA-BCY1 was subjected to immunoblotting analysis using an anti-HA antibody. Arrows on the right hand side indicate phospho-isoforms of BCY1.

6.3.9 Phosphorylation of Tpk1p does not increase in glc7-12 or glc7-13 mutants

Phosphorylation of Tpk1p increases its affinity to bind to the regulatory subunit Bcy1p. In association with Bcy1p Tpk1p is inactive. Although there have been numerous indications that Tpk1p is active, we wanted to investigate the phosphorylation status of Tpk1p in the WT and the *glc7* mutants. Any increase in the phosphorylation level of Tpk1 in the *glc7* mutants under non-permissive conditions would indicate that Glc7p regulates the phosphorylation status of Tpk1p and therefore also the activity of Tpk1p. However, no increase in the signal was observed in *glc7-13* with an antibody against the phosphorylated region of PKA in a Western analysis on HA–tagged Tpk1p. *glc7-12* even showed a reduction in phosphorylation of Tpk1p at

shift to 37º C (Fig. 6.8). We tested the same blots with an antibody directed against the HA-tag to determine whether equal amounts of protein were loaded. Altogether, these experiments suggest that Glc7p does not regulate the phosphorylation levels of Tpk1p and that Tpk1p is active in *glc7-12* and *glc7-13* under non-permissive conditions.

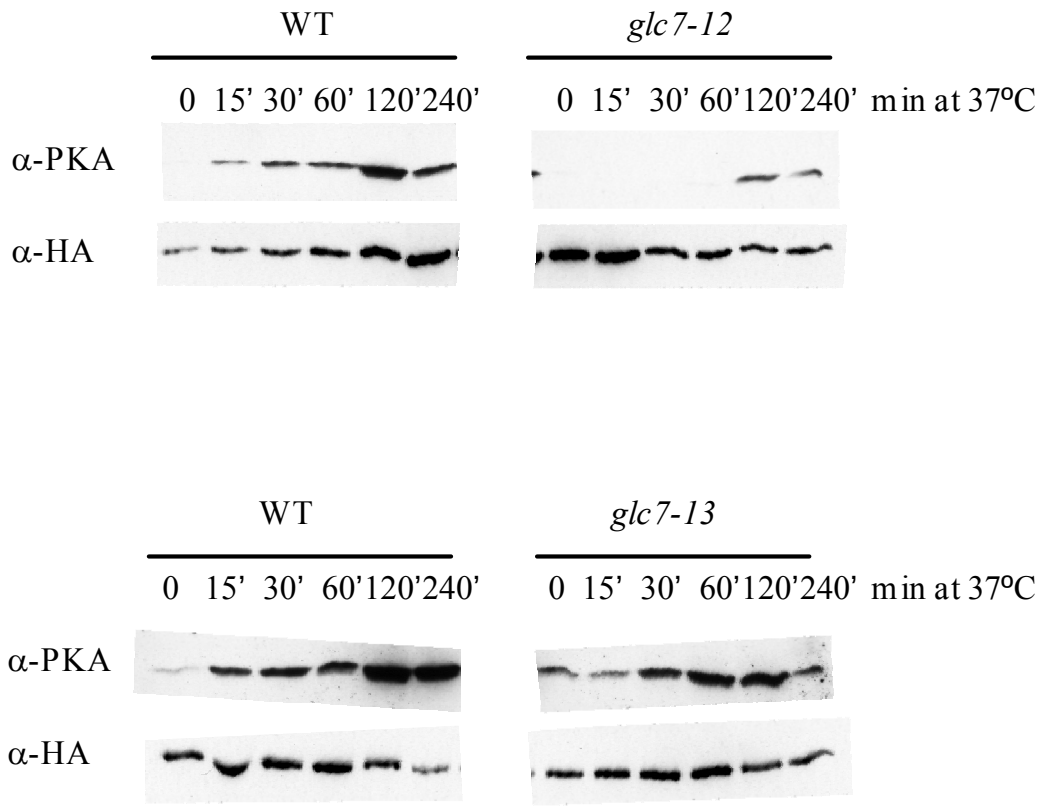


Figure 6.8. Glc7 does not regulate the phosphorylation state of TPK1
HA-tagged TPK1 was immunoprecipitated from WT, *glc7-12* (A) and *glc7-13* (B) cell extracts before and after shift to 37ºC for the indicated times. Purified HA-TPK1 was subjected to immunoblotting analysis using an anti-HA antibody or an antibody directed against the phosphorylated target region of PKA as indicated on the left.

6.3.10 Pgk1p mRNA levels are decreased in glc7-12 and glc7-13 mutants

The transcription factor Rap1p not only controls transcription of RP genes but also regulates expression of glycolytic genes such as Pgk1p. To achieve this Rap1p works in concert with different sets of transcription factors. For the transcription of RP genes Rap1p collaborates with Fhl1 and Ifh1p, whereas for the transcription of glycolytic genes Rap1 cooperates with the transcription factor Gcr1 (Drazinic et al., 1996).

To determine whether Glc7p in general influences transcription that is regulated by Rap1p, we carried out Northern analysis on total RNA extracted from the WT, *glc7-10, glc7-12*, *glc7-13*, *glc7-5* and *ssu72-2* mutant strains with a probe directed against Pgk1p after shift to the non-permissive temperature (Fig. 6.9). Indeed, we find that two of the *glc7* alleles, *glc7-12* and *glc7-13*, show a decrease in the Pgk1p mRNA. However, *glc7-10* and *glc7-5* have equal levels of the RNA over time, indicating that the downregulation of Pgk1 is allele specific. Interestingly *glc7-5* appears to influence only transcription of RP and Ribi genes but not of glycolytic genes. This experiment suggests that Glc7p is involved in transcription regulation of genes that depend on Rap1.

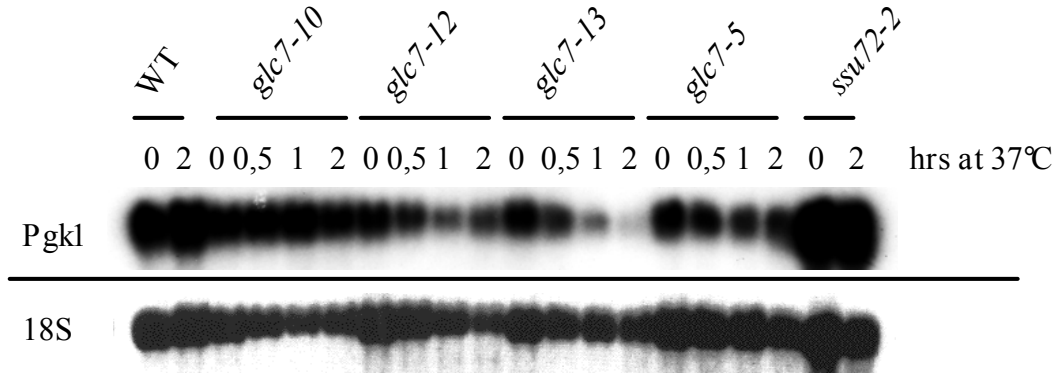


Figure 6.9. Pgk1 expression is altered in glc7 mutant strains
Northern analysis of 20µg total RNA extracted from WT, *glc7-10*, *glc7-12*, *glc7-13, glc7-5* and *ssu72-2* cells following growth at RT or after shift to 37º C for the indicated times. Filters were probed with random-primed labeled oligonucleotides directed against Pgk1. 18S served as loading control.

6.4 Discussion

Two signaling pathways control transcription of RP and Ribi genes in response to nutrient conditions: the TOR and the PKA pathways. Only recently, the transcription factors that are activated by these signaling pathways have been identified (Martin et al., 2004; Schawalder et al., 2004; Wade et al., 2004). These include Fhl1p, Ifh1p and Crf1p. We find that Glc7p is involved in the regulation of RP and Ribi gene transcription based on the finding that in *glc7* mutant strains RP and Ribi genes are downregulated as observed by both, microarray experiments and Northern analysis. Further investigation led to the observation that two downstream components of the TOR and Ras/PKA pathways, Crf1p and Yak1p, accumulated in the nucleus in *glc7* mutant strains. The mislocalization of Crf1p might be one of the reasons for the downregulation of RP genes in *glc7* mutant strains as its nuclear localization is known to repress RP gene transcription. An epistasis analysis indicated that Glc7p acts downstream of Ras. In addition, the activity of PKA was not compromised in *glc7* mutant strains at the restrictive temperature suggesting that PKA is not regulated by Glc7p. Therefore, we propose that Glc7p regulates a factor downstream of PKA or in parallel to Yak1p.

Relocalization of signaling components: a common feature in ribosome biogenesis

Localization of transcription factors to RP and/or Ribi gene promoters in response to changes in environmental conditions appears to be a common feature in the regulation of RP and Ribi gene transcription. Transcription factors that regulate transcription of RP and/or Ribi genes, such as Fhl1p, Ifh1p, Crf1p and Sfp1p have all been shown to shuttle between the nucleus and the nucleolus or the nucleus and the cytoplasm, respectively. Only for Crf1p it was clearly shown that its cellular localization is regulated via phosphorylation (Martin et al., 2004). Phosphorylation of Crf1p results in nuclear import of Crf1p and localization to RP gene promoters where it acts as a co-repressor of RP gene transcription (Martin et al., 2004). Working with the essential protein phosphatase Glc7p, we therefore considered Crf1p a prime target of Glc7p in the regulation of RP gene transcription. We find that localization of Crf1p is indeed regulated by Glc7p. Therefore, it could be possible that under optimal growth conditions Glc7p dephosphorylates nuclear Crf1p to dislocate this repressor from RP

promoters thus maintaining RP gene transcription. However, in *glc7* mutants, we also observe nuclear accumulation of Yak1p, a signaling component downstream of PKA. Yak1p is the kinase that phosphorylates Crf1p under non-favorable nutrient conditions and therefore acts upstream of Crf1p. Thus, it seems more likely that Glc7p influences localization of Crf1p via the regulation of the activity of Yak1p. We do not exclude the possibility that both, Yak1p and Crf1p are substrates for Glc7p. It was proposed that repression of RP gene transcription was mediated via the co-repressor Crf1p (Martin et al., 2004). However, the repressive property of Crf1p seems to be dependent on the strain background because deletion of Crf1p only in a TB50 strain background but not in a W303 strain background leads to a de-repression of RP gene transcription in the presence of rapamycin (Zhao et al., 2006). Presently, it is not clear what the molecular and genetic basis for the difference in Crf1p function in these two strains is (Zhao et al., 2006). In conclusion, Crf1p might not be an essential repressor of RP gene transcription.

The localization of Sfp1, a transcription factor that is involved in the regulation of both RP and Ribi genes, is also remarkably responsive to the environment (Marion et al., 2004). Under optimal growth conditions, Sfp1p is localized to the nucleus, promoting RP and Ribi gene transcription. In response to inhibition of TOR signaling or changes in nutrient availability, Sfp1p is directed to the cytoplasm. As Sfp1p is never completely excluded from the nucleus, the RP and Ribi promoters may be quite sensitive to the nuclear concentration of Sfp1p. Remarkably, the kinetics of Sfp1 relocalization to the cytoplasm matches the kinetics of RP repression (Marion et al., 2004). Sfp1 has two putative phosphorylation sites. Phosphorylation of these sites seem to be absent in rapamycin treated cells (H. Lempiainen, poster abstract). This suggests that also localization of Sfp1p is regulated by phosphorylation. We therefore plan to test whether Sfp1p is the target of Glc7p in the regulation of RP and Ribi gene transcription. Recently, it was shown that Sfp1p influences the nuclear localization of Fhl1 and Ifh1p, which are bound to RP promoters under favorable nutrient conditions (Jorgensen et al., 2004). Starvation or the absence of Sfp1p causes Fhl1p and Ifh1p to localize to nucleolar regions resulting in reduced RP gene transcription (Jorgensen et al., 2004). This observation emphasizes the important role of Sfp1p in transcription of RP genes. Mislocalization of Sfp1p in *glc7* mutant strains could therefore lead to the observed downregulation of RP genes in *glc7* mutants.

We also tested the possibility that PKA could be regulated by Glc7p. Indeed, we found that Glc7p regulates the phosphorylation status of Bcy1p, the inhibitory subunit of PKA. However, the kinetics of this regulation do not fit the kinetics of RP gene downregulation. We observed that RP gene transcription is downregulated within 15 min in glc7 mutants, whereas a difference in the phosphorylation status of Bcy1p was only visible after one hour. Thus, we concluded that regulation of RP gene transcription does not involve the dephosphorylation of Bcy1p by Glc7p. It remains to be elucidated what the function of Glc7p in regulating the phosphorylation status of Bcy1p is.

How is transcription of RP and Ribi genes regulated?

Based on recent work, a simple model evolved how RP gene transcription is regulated. In this model, Rap1p is constitutively bound to RP gene promoters and recruits Fhl1p. These two factors however are not sufficient to activate transcription of RP genes. At least the recruitment of Ifh1p by Fhl1p is required to fully activate transcription of RP genes. Several lines of evidence indicate that this model is oversimplified and that transcription of RP genes requires further factors. First, cells lacking Sfp1p do not properly regulate RP mRNA levels in response to environmental conditions (Marion et al., 2004). Secondly, residual transcription of RP genes was observed in cells without Fhl1p and Ifh1p (Rudra et al., 2005). Rapamycin represses this residual transcription suggesting that this residual transcription occurs independently of the Rap1-Fhl1-Ifh1 transcription complex (Zhao et al., 2006). These observations lead to a refined model in which inputs from both, the Sfp1p and the Rap1-Fhl1-Ifh1p complex are required for the proper activation of RP gene transcription. The localization of all of these factors to RP promoters is controlled by the TOR and the Ras/PKA pathway.

Does RP gene transcription require an repressor at all? It is possible that transcription of RP genes is repressed simply by the lack of transcriptional activators, such as, Ifh1p and Sfp1p. For example, if Ifh1p is absent only very little RP gene transcription is observed. However, transcription of RP genes has to be controlled tightly to respond to changes in the environment quickly. Therefore, it seems not appropriate that downregulation of RP gene transcription would not require a repressor. Perhaps Crf1p is only one of several redundant repressors that all respond to slightly different conditions. This could explain why a Crf1 deletion affects RP gene expression only in

certain strain backgrounds but not in others. However, the strain background-specific effect of Crf1p made it more difficult to interpret our results. Since we worked in a W303 background, it seems unlikely that the strong downregulation of RP gene transcription observed in the *glc7* mutants is only due to the nuclear accumulation of Crf1p in these strains.

Finding the real target of Glc7p in the regulation of RP gene transcription will be challenging, but important to shed light on how RP gene expression is regulated by a phosphorylation/dephosphorylation event.

7 Final Discussion

The aim of this thesis was to identify novel functions for the essential protein phosphatase Glc7p in transcription and processing of RNAP II transcripts. Glc7p is a serine/threonine protein phosphatase highly conserved in evolution that has been implicated in a quite diverse set of reactions throughout the cell.

Transcription of genes by RNAP II is a well regulated process. In order to obtain a mature, functional mRNA, the primary transcript is co-transcriptionally capped at the 5'end, spliced and 3' end processed. 3' end processing of pre-mRNAs which includes cleavage at the poly(A) site followed by polyadenylation of the upstream cleavage product requires the concerted action of the multiprotein complex CPF, CF IA and CF IB. However, the function of the different CPF subunits is not restricted to the actual cleavage and polyadenylation reaction. Many subunits of CPF are also involved in processes such as transcription elongation and termination or in coupling of 3' end formation to transcription. First indications that Glc7p could also be required for functions associated with CPF came from the finding that Glc7p co-purifies with CPF. This encouraged us to investigate a possible function of Glc7p in 3' end processing of pre-mRNAs. Indeed, we could show that Glc7p is required for the polyadenylation but not for the cleavage step of 3' end processing (Chapter 2). He and coworkers identified the CPF subunit Pta1p as the substrate of Glc7p in the regulation of polyadenylation. However, the mechanism by which phosphorylated Pta1p inhibits polyadenylation is not understood. Most likely, the phosphate-group disturbs crucial interactions between CPF subunits that are required for polyadenylation.

In addition, our investigation of Glc7p suggests that Glc7p is involved in transcription termination of snoRNAs. snoRNAs are encoded by genes that are transcribed by RNAP II. Transcription of genes terminates when RNAP II dissociates from the DNA template. Among the trans-acting factors required for snoRNA transcription termination are the Nrd1 complex, a subcomplex of CPF called APT, the nuclear exosome and the transcription elongation complex PAF. Glc7p is part of the APT subcomplex. Interestingly, *glc7* mutants show identical mutant phenotypes as mutants in all Nrd1 complex subunits, indicating that Glc7p functions in the Nrd1 complex-dependent pathway of snoRNA transcription termination. To find out how Glc7p influences snoRNA transcription termination on the molecular level it was, intriguing

to test whether the subunits of the Nrd1p complex, which are all phospho-proteins, are dephosphorylated by Glc7p. However, the phosphorylation status of none of these proteins turned out to be regulated by Glc7p. We also tested whether the phosphorylation status of Pta1p is increased in those *glc7* mutants which show a snoRNA transcription termination defect. Surprisingly, also the phosphorylation status of Pta1p was not changed in these *glc7* mutants compared to the WT, suggesting that Pta1p is not the target of Glc7p in the regulation of snoRNA transcription termination. This is interesting because it revealed that Pta1p seems to be exclusively targeted by Glc7p for the regulation of polyadenylation but not for regulation of snoRNA transcription termination. Since we have not identified the substrate of Glc7p in snoRNA transcription termination yet, it is still possible that an additional subunit of CPF is targeted by Glc7p. What complicates matters is the fact that for many CPF subunits it is not known whether they are phosphorylated or not. It would be necessary to clarify this question in order to test putative substrates of Glc7p within CPF. The question still remains which protein is regulated by Glc7p in snoRNA transcription termination. It could be one of those factors that are directly involved in snoRNA transcription termination, like a subunit of APT, one of the factors of the nuclear exosome or another still unknown factor. In addition, these factors might interact with additional factors that are needed in snoRNA transcription termination, for example for the coupling of different processes or the integration of signaling inputs. The lack of regulating the phosphorylation status of these factors could influence snoRNA transcription termination as well by disrupting important protein-protein interactions. This implicates, that the actual target of Glc7p in the regulation of snoRNA transcription termination does not necessarily have to be a factor directly involved in snoRNA transcription termination. Consequently, this complicates the search for the target in a process.

Two main problems have slowed down our efforts to elucidate the processes Glc7p is involved in and also the substrates of Glc7p.

One problem was that we did not have a conditional null allele of Glc7p to work with. Instead, we worked with temperature-sensitive (ts) *glc7* alleles. It was not predictable which ts allele has a defect in a specific process. At worst, it is possible to miss a function completely with a certain allele. We observed many allele-specific effects with the *glc7* mutant alleles we used, indicating that these alleles were not complete null alleles. As an example, only one, *glc7-13*, out of four *glc7* ts alleles showed a

polyadenylation defect. If we would not have had *glc7-13* at hand, we would have missed the fact that Glc7p is required for polyadenylation. *glc7-10* was one of the mutant alleles which mostly showed no phenotype in the processes we tested. Therefore, this allele was not useful in our analysis, although we spent much time testing all these alleles in the processes we were interested to find the one allele which showed the strongest effect.

A much more targeted approach would have been to use a conditional knock down allele, that completely abolishes any activity of the enzyme and any interactions with other proteins. With such an allele it should have been possible to test for the involvement of Glc7p directly in different processes. In addition, only one allele, the conditional knock down allele, would have to be tested and not several, as was the case for the ts alleles.

An other obstacle was that we used a candidate approach to search for the putative substrate. This means that we picked our targets based on educated guess and tested each one separately. This implies that each putative substrate had to be endogenously tagged, purified and analysed by Western analysis. This is time consuming and as described above we experienced that often the most obvious targets turned out not to be the substrates.

It might have been easier to purify entire complexes via a tagged subunit and subsequently analyze this complex for its phosphorylation state by Western analysis. Complexes like CPF are likely to harbor more than one Glc7p substrate. By analysing the entire complex candidate substrates could be revealed in one single experiment. In addition, this would also answer the question as to which of the CPF subunits are phospho-proteins at all.

Another possibility to identify substrates of Glc7p could be *in vivo* labeling with radioactive phosphorus of whole cell extracts of WT and mutant strains at the non-permissive temperature. Subsequent affinity purification of complexes and analysis of the phosphorylation status of the complex components by autoradiography could have lead to the identification of Glc7p substrates. Radioactive labeling has the advantage that differences in the phosphorylation status of putative substrates between a WT and mutant strains can be quantitated. The disadvantage of this approach is the implication of high dosages of radioactive phosphate.

Substrates identified with the approaches described should be considered with caution. This is because a change in the phosphorylation status of a protein detected

by Western or radiographic analysis does not prove that this change has been brought about by a direct dephosphorylation of this protein by Glc7p. It could be possible that Glc7p for example controls the activity of an upstream kinase that is required for the phosphorylation of this specific protein. Consequently, Glc7p would not directly regulate the phosphorylation status of this putative substrate but only indirectly influences its phosphorylation state by directly regulating the upstream kinase. The upstream kinase might not be part of the purified complex. Therefore, the substrates found by the approaches described could give misleading results and would have to be evaluated by independent approaches.

During transcription elongation specific factors assist RNAP II to efficiently transcribe genes. Transcription elongation factors either alter the chromatin structure to a transcription permissive state or change the processivity of RNAP II. These stimulating activities have to be halted to facilitate transcription termination. We found that Glc7p is involved in both processes, transcription elongation and transcription termination. Chromatin immunoprecipitation experiments demonstrated that Glc7p is predominantly localized at the promoter regions and the 3' UTRs of genes but not over the body of the gene where transcription elongation factors are mainly found. The influence of Glc7p on transcription elongation therefore might be restricted to 3' ends to alter the activities of transcription elongation factors in this region in order to facilitate transcription termination. This could be an additional function of Glc7p besides its involvement in the Nrd1p-dependent pathway of transcription termination. Glc7p could represent a dedicated transcription termination factor that acts on transcription elongation factors to halt their stimulating activities. In conclusion, Glc7p might couple transcription elongation to transcription termination. However, we can not explain why this property of Glc7p only applies to snoRNA genes. A reason for this could be that snoRNA genes including their 3'UTRs are much shorter than those of protein coding genes, implicating that the activity of RNAP II has to be controlled tighter.

Another interesting mechanism in which the function of Glc7p is involved is the transcriptional regulation of RP and Ribi genes. It could be most challenging to find the molecular function of Glc7p because many factors of the signaling pathways that control RP and Ribi gene transcription have not been identified. Nevertheless, it seems worthwhile to test one last putative substrate candidate, the transcription factor Sfp1p. Sfp1p is required for the transcription of the RP and the Ribi regulon. These

two regulons were downregulated in the *glc7-12* strain. In addition, localization of Sfp1p most likely depends on a phosphorylation event. In conclusion, the misregulation of Sfp1p could account for the downregulation of RP and Ribi genes, which we have observed in *glc7-12* and *glc7-13* mutant strains. Therefore, I am most tempted to test this hypothesis.

8 Materials and Methods

Yeast strains

Yeast strains were maintained on YPD (1% yeast extract, 2% peptone, 2% glucose) or on a selective YNB medium (0.67% yeast nitrogen base without amino acids, 2% glucose or galactose) supplemented with the appropriate nutritional ingredients. Plasmids were introduced into yeast cells using the modified lithium acetate method (Gietz et al., 1995). Loss of URA3-marked plasmids from yeast cells was accomplished by plasmid shuffling on solid medium containing 5-FOA (5-fluoro-orotic acid).

Genotypes of yeast strains used in this study were:
Glc7 strains:
PAY704-1 MATa his3 leu2 ura3 ade2 glc7::LEU2 trp1::GLC7::TRP1
PAY703-1 MATa his3 leu2 ura3 ade2 glc7::LEU2 trp1::*glc7-5*::TRP1
PAY700-4 MATa his3 leu2 ura3 ade2 glc7::LEU2 trp1::*glc7-10*::TRP1
PAY701-1 MATa his3 leu2 ura3 ade2 glc7::LEU2 trp1::*glc7-12*::TRP1
PAY702-1 MATa his3 leu2 ura3 ade2 glc7::LEU2 trp1::*glc7-13*::TRP1
Research genetics, deletion strain library:
BY4742 MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 rtf1Δ
BY4742 MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 leo1Δ
Spt5 and Spt4 strains
GHY92 MATα his4-912Δ lys2-128Δ leu2Δ1 ura3-52 spt5-242
GHY180 MATα his4-912Δ lys2-128Δ leu2Δ1 ura3-52 spt4Δ2::HIS3

Plasmids used in this study

pFA6a-CBP-TEV-protA-HIS3MX6: derived from pFA6a-HIS3MX6 (Longtine et al., 1998). Has TAP tagging cassette.
pRS416 (Sikorski and Hieter, 1989) yeast expression vector, URA3 marker
pRS426 (Sikorski and Hieter, 1989) yeast 2μ vector, URA3 marker
pKGGURA (Morillon et al., 2003b)
pKGGΔTEF (Morillon et al., 2003b)
pUGCYC1 (Birse et al., 1998b)

pDM40: CRF1 HA-tagged, 2µ, URA3, backbone YCplac33
pDM55: IFH1 HA-tagged, 2µ, URA3, backbone YCplac33
pDM77: YAK1 HA-tagged, 2µ, URA3, backbone YCplac33
pTS134: TPK1 HA-tagged, 2µ, URA3, backbone YCplac33
pTS137: BCY1 HA-tagged, 2µ, URA3, backbone YCplac33
pTS118: Ras2val19, 2µ, URA3, backbone YCplac33
M13mp18 phage vector (New England Biolabs)

Cloning experiments were performed according to established protocols (Sambrook et al., 1989).

Oligonucleotids used in this study

1) for endogenous TAP-tagging:
Sen1:
R1 5' accaatatatatgcaggtataattcctaacacttttacttcaagagaattcgagctcgtttaaac 3'
F2 5' gcttcatctagcccatttatcccaaaaaaaaagaaagcctagatcacggatccccgggttaattaa 3'
Spt5:
R1 5' ttgatttcttcttgggtgatattggttctccttttggtgagaattcgagctcgtttaaac 3'
F2 5' taagtcaaactatggtggtaacagtacatggggaggtcatcggatccccgggttaattaa 3'
Pti1:
F2 5' gcaacttatggtggaaaactttagaaaagaatatataatccggatccccggttaattaa 3'
R1 5' aattaggtgtggttttctgatacgctatggctctgattacgaattcgagctcgtttaaac 3'
Pta1:
F2 5' cttacacaagcagtgcgattcactgcttgacaggctaaaacggatccccgggttaattaa 3'
R1 5' gcctacacatgcgtatatatgatgtatgtaatggttgtgagaattcgagctcgtttaaac 3'
Pfs2:
F2 5' caggcatcgattcaaacatcaagagttccacattgccaggtttaagcatacggatccccgggttaattaa 3'
R1 5' gcctgtggggtacattccccagaaatgagagatattatcacaaataatagaattcgagctcgtttaaac 3'
Nrd1:
F2 5' gaatatgcttaaccaacagcagcagcaacaacaacaaagccggatccccgggttaattaa 3'
R1 5' ttttatgtactatgagcaaataaagggtggagtaaagatcgaattcgagctcgtttaaac 3'
Nab3:
F2 5' ctggcaataatgttcaaagtctattagatagtttagcaaaactacaaaaacggatccccgggttaattaa 3'
R1 5' agaattcaagtataatgtacaagaaatggaaaagattgaaaaaagggagtgaattcgagctcgtttaaac 3'

2) for Northern probes:

Act: 5’ atggtcggtatgggtcaaaa 3’

5’ ccaaacccaaaacagaagga 3’

Mrpl17: 5’ gggcttgctactgcaactgc 3’

5’ ccgctaaaagaaattctgttttg 3’

Ycr015c: 5’ gcaccattgctaaattaccg 3’

5’ cccctatataccaatagctgc 3’

Trs31: 5’ gcgcatctgaccaacaattc 3’

5’ gccaaatctcaaaccttccc 3’

Nrd1: 5’ gcagcaggacgacgattttc 3’

5’ ggggaggtgcgactgtac 3’

Pur5: 5’ ccgcactagactttacca 3’

5’ gggtagttctgatttttcat 3’

3) Cloning of Glc7:

Glc7NdeI: 5’ ggaactcatatggactcacaaccagttgac 3’

Glc7BamHI: 5’ tcgcggatccttatttttttctttctacccccag 3’

4) Cloning into 2µ vector:

PtiHind3 5’ gcgaagcttgcatcacatcagtgttt 3’

PtiXma1 5’ cgcccgggcacatgtgtcggagat 3’

Pta1Hind3 5’ gcgaagcttggcgtggtttccactcca 3’

Pta1Not1 5’ gggcggccgccccttcccattagata 3’

Syc1Sal1 5’ ccgtcgacccgtcaccgatgctcat 3’

Syc1BamHI 5’ cgggatccggggataaaagatataaa 3’

Coimmunoprecipitation and immunoblotting

Whole-cell extracts for SDS-PAGE analysis were prepared by glass bead lysis as described previously (Beck and Hall, 1999). Lysis buffer containing phosphatase and protease inhibitors were used as described previously (Schmelzle et al., 2004). For detection of HA-tagged proteins, a mouse anti-HA antibody (clone 12CA5) was used in a dilution of 1:5000. For detection of threonine phosphorylation an antiphosphothreonine antibody (Q7, Qiagen) was used. Signals were detected using

horseradish peroxidase-conjugated secondary goat anti-mouse antibodies and ECL reagents (Amersham Pharmacia Biotech).

Fluorescence microcopy

Fluorescence microscopy and indirect immunofluorescence on whole fixed cells were performed as described previously (Schmelzle et al., 2004). Monoclonal anti-HA (12CA5; diluted 1:500) was used as the primary antibody to detect HA-tagged proteins . Antibodies were always checked for specificity in each experiment using wild-type cells lacking the corresponding epitope.

Extract preparation and in vitro cleavage and polyadenylation assays

Extracts competent for in vitro processing were prepared following the procedure previously described (Ohnacker et al., 2000). Internally ^{32}P-labeled CYC1 RNA was produced by in vitro run-off transcription and gel purification as described previously (Minvielle-Sebastia et al., 1994; Preker et al., 1995).The cleavage and polyadenylation assays were carried out as described by (Minvielle-Sebastia et al., 1994). To restrict the assay to cleavage only, EDTA replaced MgAc and CTP replaced ATP. For reactions at elevated temperature, reaction mix and protein extracts were first preincubated separately at 37º C for 10 min, combined and assayed at 37º C for 50 min.

RNA isolation and Northern Analysis

Total RNA was isolated from 50ml yeast cultures grown to OD_{600} 0,5-1,0. RNA was isolated by hot acid phenol extraction (Collart and Oliviero, 1994). Northern analyses was carried out as described (Dichtl et al., 2002b). High molecular weight RNA was resolved on 1.2% denaturing formaldehyde/agarose gel. Random primed probe were generated using α^{32}P-dATP and the Roche random prime labeling kit. Templates were produced by PCR on yeast genomic DNA covering around 700 bp of coding sequence of the following ORF: Nrd1, Mrpl17, Pur5, Trs31, Ycr015c

Primer extension

The reverse transcription analysis was performed as described (Dichtl et al., 2002) with primers downstream of snoRNA genes as previously described (Steinmetz et al., 2001).

Expression of recombinant proteins in *E.coli*

BL21 *E.coli* cells carrying the respective plasmid were grown at 25° C in 2xYT until they reached an OD_{600} of approximately 1. Following induction by 0.5 mM IPTG incubation was continued for 6 hours. The proteins were purified at 4° C on gluthathione-Sepharose 4B as recommended (Pharmacia) and the protein was eluted with GST-elution buffer (75 mM KCl, 50 mM Tris-HCl (pH 7.9), 10% glycerol, 10 mM glutathione (reduced), 0.01% NP40, 1 mM DTT).

GST pull down experiments

In vitro translations were performed with the TNT-coupled transcription-translation system (Promega). 100 ng GST fusion protein was incubated with in 3 µl of in vitro translated [^{35}S]methionine labeled proteins for 1 hour at RT. The mixture was bound in a total volume of 1000 µl to 20 µl glutathione sepharose (Pharmacia), which was previously equilibrated in 1 ml PBS, 0.01% NP40 and 0,1mg/ml BSA. The matrix was washed 3 times with IPP150 (150 mM KCl, 20 mM Tris-Cl (pH 8,0), 0.01% NP40) and the proteins were eluted by addition of protein loading buffer and incubation at 95° C. Bound proteins were separated by 10% SDS-PAGE and visualized by autoradiography.

Droplet test

The strains were grown over night in YPD or the respective selective medium and diluted to an OD600 of 0.1, 0.001, 0.0001 and 0.00001. 5µl of each dilution was spotted on YPD plates or on selective medium and incubated at 25, 30, 33 and 37° C.

3' end labeling of total RNA

Labeling of total RNA at their 3' ends was carried out as described previously (Dichtl et al., 2002b). 6µg of total yeast RNA was labeled with α^{32}P-cordycepin (Amersham) at their 3'-ends and Rnase treated using the method of (Lingner and Keller, 1993) modified according to (Preker et al., 1995). Nuclease-resistant poly(A) tails were resolved on denaturing 8.3M urea/10% polyacrylamide gels and visualized by autoradiography.

Generation of double mutants

To generate double mutants, Glc7 strains were first transformed with plasmid pRS416 carrying GLC7 WT. Transformed strains were crossed to mutant strains GHY180, BY4742rtf1Δ, or BY4742leo1Δ. Diploid cells were sporulated and the spores were dissected. The resulting haploid strains were tested for growth on different selective media and temperatures to select the strain that had both mutations. Growth on 5FOA removed the WT plasmid from these strains and was used to investigate the effect of the double mutation on growth at different temperatures.
The genotype of the double mutants was verified by PCR and Southern analysis.

PCR-based gene modification

Plasmid pFA6a-CBP-TEV-protA-HIS3MX6 and the oligonucleotides R1 (5'-gene specific sequence-GAATTCGAGCTCGTTTAAAC-3') and F2 (5'-gene specific sequence-CGGATCCCCGGGTTAATTAA-3') were used to PCR amplify the C-terminal-TAP-tagging cassette (Longtine et al., 1998). The gene specific sequences of the forward primer were chosen to end just upstream of the stop codon preserving the reading frame of the tag, whereas those of the reverse primer were chosen to end just downstream of the stop codon. The products of six to eight PCR reactions were pooled, extracted once with phenol:chloroform:isoamyl alcohol (25:24:1), precipitated, and resuspended in 15µl of water. 5µl of this concentrated DNA was transformed into WT, *glc7-12* and *glc7-13* strains using the lithium acetate procedure. Transformants growing on SD-His medium were picked and integration of the TAP-tag cassette was confirmed by PCR.

Western analysis

Western analysis was performed according to standard protocols (Maniatis). 3% BSA was used in blocking buffer instead of milk powder in cases where anti- phospho antibodies were used.
Antibody dilutions:
8WG16 1:1000
H14 1:1000
H5 1:1000
anti-PKA 1:2000
all antibodies against CPF and CFIA subunits 1:1000

Transcriptional run-on analysis

The multi-copy plasmid pUGCYC1, in which the CYC1 gene is controlled by the strong GAL1/10 promoter, was introduced into cells of the wild-type (W303), glc7-10, *glc7-12*, *glc7-13* and glc7-5 mutant strains to achieve high levels of transcription. TRO analysis was carried out as described previously (Birse et al., 1998a). Importantly, cells were grown in SD-ura, 2% galactose medium 2 days before experiment to accommodate to this medium.

Transcriptional interference analysis

Transcriptional interference was performed as described (Morillon et al., 2003b).

Overexpression of yeast proteins

The coding regions plus 500 bp up and downstream of proteins Pta1, Pti1p and Syc1p were subcloned in 2μ vector pRS 426 (Sikorski and Hieter, 1989) and transformed into Glc7 strains. Transformed yeast strains were grown on selective medium at 25, 30, 33 and 37° C for 3 days.

cRNA Target Synthesis and GeneChip Hybridization

Following RNA isolation, the total RNA was mixed with RLT buffer. After addition of 70% ethanol, the mixture was loaded onto an RNeasy column. The total RNA was then washed and eluted in 50 μl of distilled water. Total RNA quality was monitored by loading 200 ng onto an RNA Nano 6000 Chip that was processed with the 2100 Bioanalyzer (Agilent, Palo Alto, CA). Biotin labeling of RNA was performed as described in the Expression Analysis Technical Manual (Affymetrix) with minor modifications as indicated below. To synthesize single-stranded cDNA, 5 μg of total RNA was mixed with oligo-dT and incubated with SuperScript II reverse transcriptase (Invitrogen) at 42°C for 1 h. After synthesis of the second cDNA strand by using the SuperScript double-stranded cDNA synthesis kit (Invitrogen), the material was extracted with phenol-chloroform-isoamyl alcohol, and co-precipitated with 10μg Glycogen and 0.5 volumes of 7.5 M ammonium acetate and 2.5 volumes of ethanol. cDNA was used for an *in vitro* transcription reaction by using the IVT labeling kit (Affymetrix, Santa Clara, USA) to synthesize cRNA in the presence of a

biotin-conjugated ribonucleotide analog. Approximately 100 µg of labeled cRNA from each reaction was purified by using RNeasy mini-spin columns, and roughly 300 ng was analyzed on RNA Nano 6000 Chips. The cRNA targets were incubated at 94°C for 35 min and the resulting fragments of 50–150 nucleotides were again monitored using the Bioanalyzer. All synthesis reactions were carried out using a PCR machine (T1 thermocycler; Biometra, Göttingen, Germany) to ensure the highest possible degree of temperature control.

The hybridization cocktail (200µl) containing fragmented biotin-labeled target cRNA at a final concentration of 0.05µg/µl was transferred into Yeast S98 GeneChips (Affymetrix) and incubated at 45°C on a rotator in a hybridization oven 640 (Affymetrix) for 16 h at 60 rpm. The arrays were washed and stained by using a streptavidine-phycoerythrin (SAPE) conjugate (Molecular Probes, Eugene, OR). To increase the signal strength the antibody amplification protocol was used (EukGE-WS2v5_450; Affymetrix Expression Analysis Manual). The GeneChips were processed with an Affymetrix GeneChip® Scanner 3000 (Affymetrix) by using the current default settings. DAT image files of the microarrays were generated using GeneChip Operating Software (GCOS 1.1; Affymetrix).

9 References

Adam, S.A., Nakagawa, T., Swanson, M.S., Woodruff, T.K. and Dreyfuss, G. (1986) mRNA polyadenylate-binding protein: gene isolation and sequencing and identification of a ribonucleoprotein consensus sequence. *Mol Cell Biol*, **6**, 2932-2943.

Ahn, S.H., Kim, M. and Buratowski, S. (2004) Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. *Mol Cell*, **13**, 67-76.

Alen, C., Kent, N.A., Jones, H.S., O'Sullivan, J., Aranda, A. and Proudfoot, N.J. (2002) A role for chromatin remodeling in transcriptional termination by RNA polymerase II. *Mol Cell*, **10**, 1441-1452.

Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E. and Tollervey, D. (1999) Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J*, **18**, 5399-5410.

Amrani, N., Minet, M., Wyers, F., Dufour, M.-E., Aggerbeck, L.P. and Lacroute, F. (1997a) *PCF11* encodes a third protein component of the yeast cleavage and polyadenylation factor I. *Molecular and Cellular Biology*, **17**, 1102-1109.

Anderson, C. and Tatchell, K. (2001) Hyperactive glycogen synthase mutants of Saccharomyces cerevisiae suppress the glc7-1 protein phosphatase mutant. *J Bacteriol*, **183**, 821-829.

Andrews, P.D. and Stark, M.J. (2000) Type 1 protein phosphatase is required for maintenance of cell wall integrity, morphogenesis and cell cycle progression in Saccharomyces cerevisiae. *J Cell Sci*, **113, (Pt 3)** 507-520.

Aranda, A. and Proudfoot, N. (2001) Transcriptional termination factors for RNA polymerase II in yeast. *Mol Cell*, **7**, 1003-1011.

Arigo, J.T., Carroll, K.L., Ames, J.M. and Corden, J.L. (2006a) Regulation of yeast NRD1 expression by premature transcription termination. *Mol Cell*, **21**, 641-651.

Arigo, J.T., Eyler, D.E., Carroll, K.L. and Corden, J.L. (2006b) Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. *Mol Cell*, **23**, 841-851.

Ashe, M.P., De Long, S.K. and Sachs, A.B. (2000) Glucose depletion rapidly inhibits translation initiation in yeast. *Mol Biol Cell*, **11**, 833-848.

Bachellerie, J.P., Cavaille, J. and Huttenhofer, A. (2002) The expanding snoRNA world. *Biochimie*, **84**, 775-790.

Bailis, J.M. and Roeder, G.S. (2000) Pachytene exit controlled by reversal of Mek1-dependent phosphorylation. *Cell*, **101**, 211-221.

Barabino, S.L.M., Ohnacker, M. and Keller, W. (2000) Distinct roles of two Yth1p domains in 3'-end cleavage and polyadenylation of yeast pre-mRNAs. *EMBO J*, **19**, 3778-3787.

Barbet, N.C., Schneider, U., Helliwell, S.B., Stansfield, I., Tuite, M.F. and Hall, M.N. (1996) TOR controls translation initiation and early G1 progression in yeast. *Mol Biol Cell*, **7**, 25-42.

Barford, D. (1995) Protein phosphatases. *Curr Opin Struct Biol*, **5**, 728-734.

Barford, D. (1996) Molecular mechanisms of the protein serine/threonine phosphatases. *Trends Biochem Sci*, **21**, 407-412.

Barilla, D., Lee, B.A. and Proudfoot, N.J. (2001) Cleavage/polyadenylation factor IA associates with the carboxyl- terminal domain of RNA polymerase II in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A*, **98**, 445-450.

Beck, T. and Hall, M.N. (1999) The TOR signalling pathway controls nuclear localization of nutrient-regulated transcription factors. *Nature*, **402**, 689-692.

Bell, G.I., Valenzuela, P. and Rutter, W.J. (1976) Phosphorylation of yeast RNA polymerases. *Nature*, **261**, 429-431.

Belotserkovskaya, R., Oh, S., Bondarenko, V.A., Orphanides, G., Studitsky, V.M. and Reinberg, D. (2003) FACT facilitates transcription-dependent nucleosome alteration. *Science*, **301**, 1090-1093.

Belotserkovskaya, R. and Reinberg, D. (2004) Facts about FACT and transcript elongation through chromatin. *Curr Opin Genet Dev*, **14**, 139-146.

Betz, J.L., Chang, M., Washburn, T.M., Porter, S.E., Mueller, C.L. and Jaehning, J.A. (2002) Phenotypic analysis of Paf1/RNA polymerase II complex mutations reveals connections to cell cycle regulation, protein synthesis, and lipid and nucleic acid metabolism. *Mol Genet Genomics*, **268**, 272-285.

Birse, C.E., Lee, B.A., Hansen, K. and Proudfoot, N.J. (1997) Transcriptional termination signals for RNA polymerase II in fission yeast. *The EMBO Journal*, **16**, 3633-3643.

Birse, C.E., Minvielle-Sebastia, L., Lee, B.A., Keller, W. and Proudfoot, N.J. (1998) Coupling termination of transcription to messenger RNA maturation in yeast. *Science*, **280**, 298-301.

Bloecher, A. and Tatchell, K. (2000) Dynamic localization of protein phosphatase type 1 in the mitotic cell cycle of Saccharomyces cerevisiae. *J Cell Biol*, **149**, 125-140.

Bollen, M. (2001) Combinatorial control of protein phosphatase-1. *Trends Biochem Sci*, **26**, 426-431.

Bortvin, A. and Winston, F. (1996) Evidence that Spt6p controls chromatin structure by a direct interaction with histones. *Science*, **272**, 1473-1476.

Bourgeois, C.F., Kim, Y.K., Churcher, M.J., West, M.J. and Karn, J. (2002) Spt5 cooperates with human immunodeficiency virus type 1 Tat by preventing premature RNA release at terminator sequences. *Mol Cell Biol*, **22**, 1079-1093.

Broach, J.R. and Deschenes, R.J. (1990) The function of ras genes in Saccharomyces cerevisiae. *Adv Cancer Res*, **54**, 79-139.

Bucheli, M.E. and Buratowski, S. (2005) Npl3 is an antagonist of mRNA 3' end formation by RNA polymerase II. *EMBO J*.

Buratowski, S. (2003) The CTD code. *Nat Struct Biol*, **10**, 679-680.

Buratowski, S. (2005) Connections between mRNA 3' end processing and transcription termination. *Curr Opin Cell Biol*, **17**, 257-261.

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

Callebaut, I., Moshous, D., Mornon, J.P. and de Villartay, J.P. (2002) Metallo-beta-lactamase fold within nucleic acids processing enzymes: the beta-CASP family. *Nucleic Acids Res*, **30**, 3592-3601.

Calvo, O. and Manley, J.L. (2005) The transcriptional coactivator PC4/Sub1 has multiple functions in RNA polymerase II transcription. *EMBO J*, **24**, 1009-1020.

Caponigro, G. and Parker, R. (1996) Mechanisms and control of mRNA turnover in Saccharomyces cerevisiae. *Microbiol Rev*, **60**, 233-249.

Cardenas, M.E., Cutler, N.S., Lorenz, M.C., Di Como, C.J. and Heitman, J. (1999) The TOR signaling cascade regulates gene expression in response to nutrients. *Genes Dev*, **13**, 3271-3279.

Carroll, K.L., Pradhan, D.A., Granek, J.A., Clarke, N.D. and Corden, J.L. (2004) Identification of cis elements directing termination of yeast nonpolyadenylated snoRNA transcripts. *Mol Cell Biol*, **24**, 6241-6252.

Chanfreau, G., Buckle, M. and Jacquier, A. (2000) Recognition of a conserved class of RNA tetraloops by Saccharomyces cerevisiae RNase III. *Proc Natl Acad Sci U S A*, **97**, 3142-3147.

Chanfreau, G., Elela, S.A., Ares, M., Jr. and Guthrie, C. (1997) Alternative 3'-end processing of U5 snRNA by RNase III. *Genes Dev*, **11**, 2741-2751.

Chanfreau, G., Legrain, P. and Jacquier, A. (1998) Yeast RNase III as a key processing enzyme in small nucleolar RNAs metabolism. *J Mol Biol*, **284**, 975-988.

Chen, J. and Moore, C. (1992) Separation of factors required for cleavage and polyadenylation of yeast pre-mRNA. *Mol. Cell. Biol.*, **12**, 3470-3481.

Chen, J., Zheng, X.F., Brown, E.J. and Schreiber, S.L. (1995) Identification of an 11-kDa FKBP12-rapamycin-binding domain within the 289-kDa FKBP12-rapamycin-associated protein and characterization of a critical serine residue. *Proc Natl Acad Sci U S A*, **92**, 4947-4951.

Chen, S. and Hyman, L.E. (1998) A specific RNA-protein interaction at yeast polyadenylation efficiency elements. *Nucleic Acids Research*, **26**, 4965-4974.

Cheng, H., He, X. and Moore, C. (2004) The essential WD repeat protein Swd2 has dual functions in RNA polymerase II transcription termination and lysine 4 methylation of histone H3. *Mol Cell Biol*, **24**, 2932-2943.

Cho, E.J., Kobor, M.S., Kim, M., Greenblatt, J. and Buratowski, S. (2001) Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. *Genes Dev*, **15**, 3319-3329.

Cho, E.J., Rodriguez, C.R., Takagi, T. and Buratowski, S. (1998) Allosteric interactions between capping enzyme subunits and the RNA polymerase II carboxy-terminal domain. *Genes Dev*, **12**, 3482-3487.

Cho, E.J., Takagi, T., Moore, C.R. and Buratowski, S. (1997) mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev*, **11**, 3319-3326.

Cho, H., Kim, T.K., Mancebo, H., Lane, W.S., Flores, O. and Reinberg, D. (1999) A protein phosphatase functions to recycle RNA polymerase II. *Genes Dev*, **13**, 1540-1552.

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

Cohen, P.T. (2002) Protein phosphatase 1--targeted in many directions. *J Cell Sci*, **115**, 241-256.

Colgan, D.F., Murthy, K.G., Zhao, W., Prives, C. and Manley, J.L. (1998) Inhibition of poly(A) polymerase requires p34cdc2/cyclin B phosphorylation of multiple consensus and non-consensus sites. *EMBO J*, **17**, 1053-1062.

Collart, M.A. and Oliviero, S. (eds.). (1994) *Preparation of Yeast RNA*, New York.

Conaway, J.W., Shilatifard, A., Dvir, A. and Conaway, R.C. (2000) Control of elongation by RNA polymerase II. *Trends Biochem Sci*, **25**, 375-380.

Conaway, R.C., Kong, S.E. and Conaway, J.W. (2003) TFIIS and GreB: two like-minded transcription elongation factors with sticky fingers. *Cell*, **114**, 272-274.

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

Connelly, S. and Manley, J.L. (1989a) A CCAAT box sequence in the adenovirus major late promoter functions as part of an RNA polymerase II termination signal. *Cell*, **57**, 561-571.

Connelly, S. and Manley, J.L. (1989b) RNA polymerase II transcription termination is mediated specifically by protein binding to a CCAAT box sequence. *Mol Cell Biol*, **9**, 5254-5259.

Conrad, N.K., Wilson, S.M., Steinmetz, E.J., Patturajan, M., Brow, D.A., Swanson, M.S. and Corden, J.L. (2000) A yeast heterogeneous nuclear ribonucleoprotein complex associated with RNA polymerase II. *Genetics*, **154**, 557-571.

Costa, P.J. and Arndt, K.M. (2000) Synthetic lethal interactions suggest a role for the Saccharomyces cerevisiae Rtf1 protein in transcription elongation. *Genetics*, **156**, 535-547.

Crespo, J.L. and Hall, M.N. (2002) Elucidating TOR signaling and rapamycin action: lessons from Saccharomyces cerevisiae. *Microbiol Mol Biol Rev*, **66**, 579-591, table of contents.

Cui, D.Y., Brown, C.R. and Chiang, H.L. (2004) The type 1 phosphatase Reg1p-Glc7p is required for the glucose-induced degradation of fructose-1,6-bisphosphatase in the vacuole. *J Biol Chem*, **279**, 9713-9724.

Dahmus, M.E. (1996) Reversible phosphorylation of the C-terminal domain of RNA polymerase II. *J Biol Chem*, **271**, 19009-19012.

Davie, J.K. and Kane, C.M. (2000) Genetic interactions between TFIIS and the Swi-Snf chromatin-remodeling complex. *Mol Cell Biol*, **20**, 5960-5973.

Depaoli-Roach, A.A., Park, I.K., Cerovsky, V., Csortos, C., Durbin, S.D., Kuntz, M.J., Sitikov, A., Tang, P.M., Verin, A. and Zolnierowicz, S. (1994) Serine/threonine protein phosphatases in the control of cell function. *Adv Enzyme Regul*, **34**, 199-224.

Dheur, S., Vo le, T.A., Voisinet-Hakil, F., Minet, M., Schmitter, J.M., Lacroute, F., Wyers, F. and Minvielle-Sebastia, L. (2003) Pti1p and Ref2p found in association with the mRNA 3' end formation complex direct snoRNA maturation. *EMBO J*, **22**, 2831-2840.

Dichtl, B., Aasland, R. and Keller, W. (2004) Functions for S. cerevisiae Swd2p in 3' end formation of specific mRNAs and snoRNAs and global histone 3 lysine 4 methylation. *RNA*, **10**, 965-977.

Dichtl, B., Blank, D., Ohnacker, M., Friedlein, A., Roeder, D., Langen, H. and Keller, W. (2002a) A role for SSU72 in balancing RNA polymerase II transcription elongation and termination. *Mol Cell*, **10**, 1139-1150.

Dichtl, B., Blank, D., Sadowski, M., Hübner, W., Weiser, S. and Keller, W. (2002b) Yhh1p/Cft1p directly links poly(A) site recognition and RNA polymerase II transcription termination. *EMBO J*, **21**, 4125-4135.

Dichtl, B. and Keller, W. (2001) Recognition of polyadenylation sites in yeast pre-mRNAs by cleavage and polyadenylation factor. *EMBO J.*, **20**, 3197-3209.

Dombek, K.M., Kacherovsky, N. and Young, E.T. (2004) The Reg1-interacting proteins, Bmh1, Bmh2, Ssb1, and Ssb2, have roles in maintaining glucose repression in Saccharomyces cerevisiae. *J Biol Chem*, **279**, 39165-39174.

Drazinic, C.M., Smerage, J.B., Lopez, M.C. and Baker, H.V. (1996) Activation mechanism of the multifunctional transcription factor repressor-activator protein 1 (Rap1p). *Mol Cell Biol*, **16**, 3187-3196.

Dye, M.J. and Proudfoot, N.J. (1999) Terminal exon definition occurs cotranscriptionally and promotes termination of RNA polymerase II. *Mol Cell*, **3**, 371-378.

Edwalds-Gilbert, G., Prescott, J. and Falck-Pedersen, E. (1993) 3' RNA processing efficiency plays a primary role in generating termination-competent RNA polymerase II elongation complexes. *Mol Cell Biol*, **13**, 3472-3480.

Egecioglu, D.E., Henras, A.K. and Chanfreau, G.F. (2006) Contributions of Trf4p- and Trf5p-dependent polyadenylation to the processing and degradative functions of the yeast nuclear exosome. *Rna*, **12**, 26-32.

Fatica, A., Morlando, M. and Bozzoni, I. (2000) Yeast snoRNA accumulation relies on a cleavage- dependent/polyadenylation-independent 3'-processing apparatus. *EMBO J*, **19**, 6218-6229.

Fatica, A. and Tollervey, D. (2002) Making ribosomes. *Curr Opin Cell Biol*, **14**, 313-318.

Feng, Z.H., Wilson, S.E., Peng, Z.Y., Schlender, K.K., Reimann, E.M. and Trumbly, R.J. (1991) The yeast GLC7 gene required for glycogen accumulation encodes a type 1 protein phosphatase. *J Biol Chem*, **266**, 23796-23801.

Fong, N. and Bentley, D.L. (2001) Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD. *Genes Dev*, **15**, 1783-1795.

Fujinaga, K., Irwin, D., Huang, Y., Taube, R., Kurosu, T. and Peterlin, B.M. (2004) Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. *Mol Cell Biol*, **24**, 787-795.

Ganem, C., Devaux, F., Torchet, C., Jacq, C., Quevillon-Cheruel, S., Labesse, G., Facca, C. and Faye, G. (2003) Ssu72 is a phosphatase essential for transcription termination of snoRNAs and specific mRNAs in yeast. *EMBO J*, **22**, 1588-1598.

Garcia-Gimeno, M.A., Munoz, I., Arino, J. and Sanz, P. (2003) Molecular characterization of Ypi1, a novel Saccharomyces cerevisiae type 1 protein phosphatase inhibitor. *J Biol Chem*, **278**, 47744-47752.

Garrett, S. and Broach, J. (1989) Loss of Ras activity in Saccharomyces cerevisiae is suppressed by disruptions of a new kinase gene, YAKI, whose product may act downstream of the cAMP-dependent protein kinase. *Genes Dev*, **3**, 1336-1348.

Gavin, A.C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dumpelfeld, B., Edelmann, A., Heurtier, M.A., Hoffman, V., Hoefert, C., Klein, K., Hudak, M., Michon, A.M., Schelder, M., Schirle, M., Remor, M., Rudi, T., Hooper, S., Bauer, A., Bouwmeester, T., Casari, G., Drewes, G., Neubauer, G., Rick, J.M., Kuster, B., Bork, P., Russell, R.B. and Superti-Furga, G. (2006) Proteome survey reveals modularity of the yeast cell machinery. *Nature*, **440**, 631-636.

Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., Remor, M., Hofert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M.A., Copley, R.R., Edelmann, A., Querfurth, E., Rybin, V.,

Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G. and Superti-Furga, G. (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature*, **415**, 141-147.

Gietz, R.D., Schiestl, R.H., Willems, A.R. and Woods, R.A. (1995) Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast*, **11**, 355-360.

Gilbert, W. and Guthrie, C. (2004) The Glc7p nuclear phosphatase promotes mRNA export by facilitating association of Mex67p with mRNA. *Mol Cell*, **13**, 201-212.

Gonzalez, C.I., Ruiz-Echevarria, M.J., Vasudevan, S., Henry, M.F. and Peltz, S.W. (2000) The yeast hnRNP-like protein Hrp1/Nab4 marks a transcript for nonsense-mediated mRNA decay. *Mol Cell*, **5**, 489-499.

Gorner, W., Durchschlag, E., Martinez-Pastor, M.T., Estruch, F., Ammerer, G., Hamilton, B., Ruis, H. and Schuller, C. (1998) Nuclear localization of the C2H2 zinc finger protein Msn2p is regulated by stress and protein kinase A activity. *Genes Dev*, **12**, 586-597.

Graber, J.H., Cantor, C.R., Mohr, S.C. and Smith, T.F. (1999a) In silico detection of control signals: mRNA 3'-end-processing sequences in diverse species. *Proc Natl Acad Sci U S A*, **96**, 14055-14060.

Graber, J.H., Cantor, C.R., Mohr, S.C. and Smith, T.F. (1999b) Genomic detection of new yeast pre-mRNA 3'-end-processing signals. *Nucleic Acids Res*, **27**, 888-894.

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

Greger, I.H., Aranda, A. and Proudfoot, N. (2000) Balancing transcriptional interference and initiation on the GAL7 promoter of Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A*, **97**, 8415-8420.

Greger, I.H. and Proudfoot, N.J. (1998) Poly(A) signals control both transcriptional termination and initiation between the tandem GAL10 and GAL7 genes of Saccharomyces cerevisiae. *EMBO J*, **17**, 4771-4779.

Griffioen, G., Anghileri, P., Imre, E., Baroni, M.D. and Ruis, H. (2000) Nutritional control of nucleocytoplasmic localization of cAMP-dependent protein kinase catalytic and regulatory subunits in Saccharomyces cerevisiae. *J Biol Chem*, **275**, 1449-1456.

Griffioen, G., Branduardi, P., Ballarini, A., Anghileri, P., Norbeck, J., Baroni, M.D. and Ruis, H. (2001) Nucleocytoplasmic distribution of budding yeast protein kinase A regulatory subunit Bcy1 requires Zds1 and is regulated by Yak1-dependent phosphorylation of its targeting domain. *Mol Cell Biol*, **21**, 511-523.

Griffioen, G., Swinnen, S. and Thevelein, J.M. (2003) Feedback inhibition on cell wall integrity signaling by Zds1 involves Gsk3 phosphorylation of a cAMP-dependent protein kinase regulatory subunit. *J Biol Chem*, **278**, 23460-23471.

Grigull, J., Mnaimneh, S., Pootoolal, J., Robinson, M.D. and Hughes, T.R. (2004) Genome-wide analysis of mRNA stability using transcription inhibitors and microarrays reveals posttranscriptional control of ribosome biogenesis factors. *Mol Cell Biol*, **24**, 5534-5547.

Gross, S. and Moore, C. (2001a) Five subunits are required for reconstitution of the cleavage and polyadenylation activities of Saccharomyces cerevisiae cleavage factor I. *Proc Natl Acad Sci U S A*, **98**, 6080-6085.

Gross, S. and Moore, C.L. (2001b) Rna15 interaction with the A-rich yeast polyadenylation signal is an essential step in mRNA 3'-end formation. *Mol Cell Biol*, **21**, 8045-8055.

Guo, Z., Russo, P., Yun, D.-F., Butler, J.S. and Sherman, F. (1995) Redundant 3' end-forming signals for the yeast *CYC1 mRNA*. *Proceedings of the National Academy of Sciences USA*, **92**, 4211-4214.

Guo, Z. and Sherman, F. (1996) 3'-end-forming signals of yeast mRNA. *Trends in Biochemical Sciences*, **21**, 477-481.

Hampsey, M. (1997) A review of phenotypes in Saccharomyces cerevisiae. *Yeast*, **13**, 1099-1133.

Hartzog, G.A., Speer, J.L. and Lindstrom, D.L. (2002) Transcript elongation on a nucleoprotein template. *Biochim Biophys Acta*, **1577**, 276-286.

Hartzog, G.A., Wada, T., Handa, H. and Winston, F. (1998) Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in Saccharomyces cerevisiae. *Genes Dev*, **12**, 357-369.

He, X., Khan, A.U., Cheng, H., Pappas, D.L., Jr., Hampsey, M. and Moore, C.L. (2003) Functional interactions between the transcription and mRNA 3' end processing machineries mediated by Ssu72 and Sub1. *Genes Dev*, **17**, 1030-1042.

He, X. and Moore, C. (2005) Regulation of yeast mRNA 3' end processing by phosphorylation. *Mol Cell*, **19**, 619-629.

Heidmann, S., Schwindewolf, C., Stumpf, G. and Domdey, H. (1994) Flexibility and interchangeability of polyadenylation signals in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, **14**, 4633-4642.

Helliwell, S.B., Wagner, P., Kunz, J., Deuter-Reinhard, M., Henriquez, R. and Hall, M.N. (1994) TOR1 and TOR2 are structurally and functionally similar but not identical phosphatidylinositol kinase homologues in yeast. *Mol Biol Cell*, **5**, 105-118.

Henry, M., Borland, C.Z., Bossie, M. and Silver, P.A. (1996) Potential RNA binding proteins in Saccharomyces cerevisiae identified as suppressors of temperature-sensitive mutations in NPL3. *Genetics*, **142**, 103-115.

Hentze, M.W. and Kulozik, A.E. (1999) A perfect message: RNA surveillance and nonsense-mediated decay. *Cell*, **96**, 307-310.

Hirose, Y. and Manley, J.L. (1998) RNA polymerase II is an essential mRNA polyadenylation factor. *Nature (London)*, **395**, 93-96.

Hirose, Y. and Manley, J.L. (2000) RNA polymerase II and the integration of nuclear events. *Genes Dev*, **14**, 1415-1429.

Ho, C.K., Sriskanda, V., McCracken, S., Bentley, D., Schwer, B. and Shuman, S. (1998) The guanylyltransferase domain of mammalian mRNA capping enzyme binds to the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J Biol Chem*, **273**, 9577-9585.

Ho, Y., Gruhler, A., Heilbut, A., Bader, G.D., Moore, L., Adams, S.L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., Yang, L., Wolting, C., Donaldson, I., Schandorff, S., Shewnarane, J., Vo, M., Taggart, J., Goudreault, M., Muskat, B., Alfarano, C., Dewar, D., Lin, Z., Michalickova, K., Willems, A.R., Sassi, H., Nielsen, P.A., Rasmussen, K.J., Andersen, J.R., Johansen, L.E., Hansen, L.H., Jespersen, H., Podtelejnikov, A., Nielsen, E., Crawford, J., Poulsen, V., Sorensen, B.D., Matthiesen, J., Hendrickson, R.C., Gleeson, F., Pawson, T., Moran, M.F., Durocher, D., Mann, M., Hogue, C.W., Figeys, D. and Tyers, M.

(2002) Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. *Nature*, **415**, 180-183.

Hong, G., Trumbly, R.J., Reimann, E.M. and Schlender, K.K. (2000) Sds22p is a subunit of a stable isolatable form of protein phosphatase 1 (Glc7p) from Saccharomyces cerevisiae. *Arch Biochem Biophys*, **376**, 288-298.

Howard, S.C., Hester, A. and Herman, P.K. (2003) The Ras/PKA signaling pathway may control RNA polymerase II elongation via the Spt4p/Spt5p complex in Saccharomyces cerevisiae. *Genetics*, **165**, 1059-1070.

Howe, K. (2002) RNA polymerase II conducts a symphony of pre-mRNA processing activities. *Biochim Biophys Acta*, **1577**, 308.

Hsu, J.Y., Sun, Z.W., Li, X., Reuben, M., Tatchell, K., Bishop, D.K., Grushcow, J.M., Brame, C.J., Caldwell, J.A., Hunt, D.F., Lin, R., Smith, M.M. and Allis, C.D. (2000) Mitotic phosphorylation of histone H3 is governed by Ipl1/aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes. *Cell*, **102**, 279-291.

Ito, T., Chiba, T. and Yoshida, M. (2001) Exploring the protein interactome using comprehensive two-hybrid projects. *Trends Biotechnol*, **19**, S23-27.

Ito, T., Tashiro, K., Muta, S., Ozawa, R., Chiba, T., Nishizawa, M., Yamamoto, K., Kuhara, S. and Sakaki, Y. (2000) Toward a protein-protein interaction map of the budding yeast: A comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins. *Proc Natl Acad Sci U S A*, **97**, 1143-1147.

Ivanov, D., Kwak, Y.T., Guo, J. and Gaynor, R.B. (2000) Domains in the SPT5 protein that modulate its transcriptional regulatory properties. *Mol Cell Biol*, **20**, 2970-2983.

Jones, J.C., Phatnani, H.P., Haystead, T.A., MacDonald, J.A., Alam, S.M. and Greenleaf, A.L. (2004) C-terminal repeat domain kinase I phosphorylates Ser2 and Ser5 of RNA polymerase II C-terminal domain repeats. *J Biol Chem*, **279**, 24957-24964.

Jorgensen, P., Nishikawa, J.L., Breitkreutz, B.J. and Tyers, M. (2002) Systematic identification of pathways that couple cell growth and division in yeast. *Science*, **297**, 395-400.

Jorgensen, P., Rupes, I., Sharom, J.R., Schneper, L., Broach, J.R. and Tyers, M. (2004) A dynamic transcriptional network communicates growth potential to ribosome synthesis and critical cell size. *Genes Dev*, **18**, 2491-2505.

Kamada, K., Roeder, R.G. and Burley, S.K. (2003) Molecular mechanism of recruitment of TFIIF- associating RNA polymerase C-terminal domain phosphatase (FCP1) by transcription factor IIF. *Proc Natl Acad Sci U S A*, **100**, 2296-2299.

Kaplan, C.D., Holland, M.J. and Winston, F. (2005) Interaction between transcription elongation factors and mRNA 3'-end formation at the Saccharomyces cerevisiae GAL10-GAL7 locus. *J Biol Chem*, **280**, 913-922.

Keller, W. and Minvielle-Sebastia, L. (1997) A comparison of mammalian and yeast pre-mRNA 3'-end processing. *Current Opinion in Cell Biology*, **9**, 329-336.

Kelley, D.E., Stokes, D.G. and Perry, R.P. (1999) CHD1 interacts with SSRP1 and depends on both its chromodomain and its ATPase/helicase-like domain for proper association with chromatin. *Chromosoma*, **108**, 10-25.

Kessler, M.M., Henry, M.F., Shen, E., Zhao, J., Gross, S., Silver, P.A. and Moore, C.L. (1997) Hrp1, a sequence-specific RNA-binding protein that shuttles

between the nucleus and the cytoplasm, is required for mRNA 3'-end formation in yeast. *Genes and Development*, **11**, 2545-2556.

Kessler, M.M., Zhao, J. and Moore, C.L. (1996) Purification of the *Saccharomyces cerevisiae* cleavage/polyadenylation factor I. *J. Biol. Chem.*, **271**, 27167-27175.

Kettenberger, H., Armache, K.J. and Cramer, P. (2003) Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. *Cell*, **114**, 347-357.

Khorasanizadeh, S. (2004) The nucleosome: from genomic organization to genomic regulation. *Cell*, **116**, 259-272.

Kim, J.B. and Sharp, P.A. (2001) Positive transcription elongation factor B phosphorylates hSPT5 and RNA polymerase II carboxyl-terminal domain independently of cyclin-dependent kinase-activating kinase. *J Biol Chem*, **276**, 12317-12323.

Kim, M., Ahn, S.H., Krogan, N.J., Greenblatt, J.F. and Buratowski, S. (2004a) Transitions in RNA polymerase II elongation complexes at the 3' ends of genes. *EMBO J*, **23**, 354-364.

Kim, M., Krogan, N.J., Vasiljeva, L., Rando, O.J., Nedea, E., Greenblatt, J.F. and Buratowski, S. (2004b) The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature*, **432**, 517-522.

Kiss, T. (2001) Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. *EMBO J*, **20**, 3617-3622.

Kiss, T. (2002) Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. *Cell*, **109**, 145-148.

Kizer, K.O., Phatnani, H.P., Shibata, Y., Hall, H., Greenleaf, A.L. and Strahl, B.D. (2005) A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. *Mol Cell Biol*, **25**, 3305-3316.

Klein, C. and Struhl, K. (1994) Protein kinase A mediates growth-regulated expression of yeast ribosomal protein genes by modulating RAP1 transcriptional activity. *Mol Cell Biol*, **14**, 1920-1928.

Komarnitsky, P., Cho, E.J. and Buratowski, S. (2000) Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev*, **14**, 2452-2460.

Kops, O., Zhou, X.Z. and Lu, K.P. (2002) Pin1 modulates the dephosphorylation of the RNA polymerase II C-terminal domain by yeast Fcp1. *FEBS Lett*, **513**, 305-311.

Kramer, A. (1996) The structure and function of proteins involved in mammalian pre-mRNA splicing. *Annu Rev Biochem*, **65**, 367-409.

Krishnamurthy, S., He, X., Reyes-Reyes, M., Moore, C. and Hampsey, M. (2004) Ssu72 is an RNA polymerase II CTD phosphatase. *Mol Cell*, **14**, 387-394.

Krogan, N.J., Dover, J., Wood, A., Schneider, J., Heidt, J., Boateng, M.A., Dean, K., Ryan, O.W., Golshani, A., Johnston, M., Greenblatt, J.F. and Shilatifard, A. (2003a) The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. *Mol Cell*, **11**, 721-729.

Krogan, N.J., Kim, M., Ahn, S.H., Zhong, G., Kobor, M.S., Cagney, G., Emili, A., Shilatifard, A., Buratowski, S. and Greenblatt, J.F. (2002) RNA polymerase II elongation factors of Saccharomyces cerevisiae: a targeted proteomics approach. *Mol Cell Biol*, **22**, 6979-6992.

Krogan, N.J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D.P., Beattie, B.K., Emili, A., Boone, C., Shilatifard, A., Buratowski, S. and Greenblatt, J. (2003b) Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. *Mol Cell Biol*, **23**, 4207-4218.

Kyburz, A., Sadowski, M., Dichtl, B. and Keller, W. (2003) The role of the yeast cleavage and polyadenylation factor subunit Ydh1p/Cft2p in pre-mRNA 3'-end formation. *Nucleic Acids Res*, **31**, 3936-3945.

Lachner, M., O'Sullivan, R.J. and Jenuwein, T. (2003) An epigenetic road map for histone lysine methylation. *J Cell Sci*, **116**, 2117-2124.

Lai, C.C., Boguski, M., Broek, D. and Powers, S. (1993) Influence of guanine nucleotides on complex formation between Ras and CDC25 proteins. *Mol Cell Biol*, **13**, 1345-1352.

Lee, C.Y., Lee, A. and Chanfreau, G. (2003) The roles of endonucleolytic cleavage and exonucleolytic digestion in the 5'-end processing of S. cerevisiae box C/D snoRNAs. *Rna*, **9**, 1362-1370.

Lee, T.I., Rinaldi, N.J., Robert, F., Odom, D.T., Bar-Joseph, Z., Gerber, G.K., Hannett, N.M., Harbison, C.T., Thompson, C.M., Simon, I., Zeitlinger, J., Jennings, E.G., Murray, H.L., Gordon, D.B., Ren, B., Wyrick, J.J., Tagne, J.B., Volkert, T.L., Fraenkel, E., Gifford, D.K. and Young, R.A. (2002) Transcriptional regulatory networks in Saccharomyces cerevisiae. *Science*, **298**, 799-804.

Li, M., Phatnani, H.P., Guan, Z., Sage, H., Greenleaf, A.L. and Zhou, P. (2005) Solution structure of the Set2-Rpb1 interacting domain of human Set2 and its interaction with the hyperphosphorylated C-terminal domain of Rpb1. *Proc Natl Acad Sci U S A*, **102**, 17636-17641.

Licatalosi, D.D., Geiger, G., Minet, M., Schroeder, S., Cilli, K., McNeil, J.B. and Bentley, D.L. (2002) Functional Interaction of Yeast Pre-mRNA 3' End Processing Factors with RNA Polymerase II. *Mol Cell*, **9**, 1101-1111.

Lin, P.S., Dubois, M.F. and Dahmus, M.E. (2002) TFIIF-associating carboxyl-terminal domain phosphatase dephosphorylates phosphoserines 2 and 5 of RNA polymerase II. *J Biol Chem*, **277**, 45949-45956.

Lindstrom, D.L. and Hartzog, G.A. (2001) Genetic interactions of Spt4-Spt5 and TFIIS with the RNA polymerase II CTD and CTD modifying enzymes in Saccharomyces cerevisiae. *Genetics*, **159**, 487-497.

Lindstrom, D.L., Squazzo, S.L., Muster, N., Burckin, T.A., Wachter, K.C., Emigh, C.A., McCleery, J.A., Yates, J.R., 3rd and Hartzog, G.A. (2003) Dual roles for Spt5 in pre-mRNA processing and transcription elongation revealed by identification of Spt5-associated proteins. *Mol Cell Biol*, **23**, 1368-1378.

Lingner, J. and Keller, W. (1993) 3'-end labeling of RNA with recombinant yeast poly(A) polymerase. *Nucleic Acids Res*, **21**, 2917-2920.

Lis, J. (1998) Promoter-associated pausing in promoter architecture and postinitiation transcriptional regulation. *Cold Spring Harb Symp Quant Biol*, **63**, 347-356.

Logan, J., Falck-Pedersen, E., Darnell, J.E., Jr. and Shenk, T. (1987) A poly(A) addition site and a downstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse beta maj-globin gene. *Proc Natl Acad Sci U S A*, **84**, 8306-8310.

Long, R.M., Elliott, D.J., Stutz, F., Rosbash, M. and Singer, R.H. (1995) Spatial consequences of defective processing of specific yeast mRNAs revealed by fluorescent in situ hybridization. *Rna*, **1**, 1071-1078.

Longtine, M.S., McKenzie, A., 3rd, Demarini, D.J., Shah, N.G., Wach, A., Brachat, A., Philippsen, P. and Pringle, J.R. (1998) Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. *Yeast*, **14**, 953-961.

Luo, W. and Bentley, D. (2004) A ribonucleolytic rat torpedoes RNA polymerase II. *Cell*, **119**, 911-914.

Luo, W., Johnson, A.W. and Bentley, D.L. (2006) The role of Rat1 in coupling mRNA 3'-end processing to transcription termination: implications for a unified allosteric-torpedo model. *Genes Dev*, **20**, 954-965.

Lusser, A. and Kadonaga, J.T. (2003) Chromatin remodeling by ATP-dependent molecular machines. *Bioessays*, **25**, 1192-1200.

MacKelvie, S.H., Andrews, P.D. and Stark, M.J. (1995) The Saccharomyces cerevisiae gene SDS22 encodes a potential regulator of the mitotic function of yeast type 1 protein phosphatase. *Mol Cell Biol*, **15**, 3777-3785.

Mandal, S.S., Cho, H., Kim, S., Cabane, K. and Reinberg, D. (2002) FCP1, a phosphatase specific for the heptapeptide repeat of the largest subunit of RNA polymerase II, stimulates transcription elongation. *Mol Cell Biol*, **22**, 7543-7552.

Mandart, E. and Parker, R. (1995) Effects of mutations in the Saccharomyces cerevisiae RNA14, RNA15, and PAP1 genes on polyadenylation in vivo. *Mol Cell Biol*, **15**, 6979-6986.

Maniatis, T. and Reed, R. (2002) An extensive network of coupling among gene expression machines. *Nature*, **416**, 499-506.

Manley, J.L. (2002) Nuclear coupling: RNA processing reaches back to transcription. *Nat Struct Biol*, **9**, 790-791.

Marion, R.M., Regev, A., Segal, E., Barash, Y., Koller, D., Friedman, N. and O'Shea, E.K. (2004) Sfp1 is a stress- and nutrient-sensitive regulator of ribosomal protein gene expression. *Proc Natl Acad Sci U S A*, **101**, 14315-14322.

Marshall, N.F., Peng, J., Xie, Z. and Price, D.H. (1996) Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. *J Biol Chem*, **271**, 27176-27183.

Martin, D.E., Soulard, A. and Hall, M.N. (2004) TOR regulates ribosomal protein gene expression via PKA and the Forkhead transcription factor FHL1. *Cell*, **119**, 969-979.

Mayordomo, I., Regelmann, J., Horak, J. and Sanz, P. (2003) Saccharomyces cerevisiae 14-3-3 proteins Bmh1 and Bmh2 participate in the process of catabolite inactivation of maltose permease. *FEBS Lett*, **544**, 160-164.

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

McNeil, J.B., Agah, H. and Bentley, D. (1998) Activated transcription independent of the RNA polymerase II holoenzyme in budding yeast. *Genes Dev*, **12**, 2510-2521.

Meinhart, A., Kamenski, T., Hoeppner, S., Baumli, S. and Cramer, P. (2005) A structural perspective of CTD function. *Genes Dev*, **19**, 1401-1415.

Minvielle-Sebastia, L., Beyer, K., Krecic, A.M., Hector, R.E., Swanson, M.S. and Keller, W. (1998) Control of cleavage site selection during mRNA 3'-end formation by a yeast hnRNP. *The EMBO Journal*, **17**, 7454-7468.

Minvielle-Sebastia, L., Preker, P.J. and Keller, W. (1994) RNA14 and RNA15 proteins as components of a yeast pre-mRNA 3'-end processing factor. *Science*, **266**, 1702-1705.

Minvielle-Sebastia, L., Preker, P.J., Wiederkehr, T., Strahm, Y. and Keller, W. (1997) The major yeast poly(A)-binding protein is associated with cleavage factor IA and functions in pre-messenger RNA 3'-end formation. *Proceedings of the National Academy of Sciences USA*, **94**, 7897-7902.

Minvielle-Sebastia, L., Winsor, B., Bonneaud, N. and Lacroute, F. (1991) Mutations in the yeast *RNA14* and *RNA15* genes result in an abnormal mRNA decay rate; Sequence analysis reveals an RNA-binding domain in the RNA15 protein. *Mol. Cell. Biol.*, **11**, 3075-3087.

Mizrahi, N. and Moore, C. (2000) Posttranslational phosphorylation and ubiquitination of the Saccharomyces cerevisiae Poly(A) polymerase at the S/G(2) stage of the cell cycle. *Mol Cell Biol*, **20**, 2794-2802.

Morillon, A., Karabetsou, N., O'Sullivan, J., Kent, N., Proudfoot, N. and Mellor, J. (2003a) Isw1 chromatin remodeling ATPase coordinates transcription elongation and termination by RNA polymerase II. *Cell*, **115**, 425-435.

Morillon, A., O'Sullivan, J., Azad, A., Proudfoot, N. and Mellor, J. (2003b) Regulation of elongating RNA polymerase II by forkhead transcription factors in yeast. *Science*, **300**, 492-495.

Moriya, H., Shimizu-Yoshida, Y., Omori, A., Iwashita, S., Katoh, M. and Sakai, A. (2001) Yak1p, a DYRK family kinase, translocates to the nucleus and phosphorylates yeast Pop2p in response to a glucose signal. *Genes Dev*, **15**, 1217-1228.

Morlando, M., Ballarino, M., Greco, P., Caffarelli, E., Dichtl, B. and Bozzoni, I. (2004) Coupling between snoRNP assembly and 3' processing controls box C/D snoRNA biosynthesis in yeast. *EMBO J*, **23**, 2392-2401.

Morlando, M., Greco, P., Dichtl, B., Fatica, A., Keller, W. and Bozzoni, I. (2002) Functional analysis of yeast snoRNA and snRNA 3'-end formation mediated by uncoupling of cleavage and polyadenylation. *Mol Cell Biol*, **22**, 1379-1389.

Morris, D.P., Michelotti, G.A. and Schwinn, D.A. (2005) Evidence that phosphorylation of the RNA polymerase II carboxyl-terminal repeats is similar in yeast and humans. *J Biol Chem*, **280**, 31368-31377.

Morse, R.H. (2000) RAP, RAP, open up! New wrinkles for RAP1 in yeast. *Trends Genet*, **16**, 51-53.

Mueller, C.L. and Jaehning, J.A. (2002) Ctr9, Rtf1, and Leo1 are components of the Paf1/RNA polymerase II complex. *Mol Cell Biol*, **22**, 1971-1980.

Mueller, C.L., Porter, S.E., Hoffman, M.G. and Jaehning, J.A. (2004) The Paf1 complex has functions independent of actively transcribing RNA polymerase II. *Mol Cell*, **14**, 447-456.

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

Narlikar, G.J., Fan, H.Y. and Kingston, R.E. (2002) Cooperation between complexes that regulate chromatin structure and transcription. *Cell*, **108**, 475-487.

Nedea, E., He, X., Kim, M., Pootoolal, J., Zhong, G., Canadien, V., Hughes, T., Buratowski, S., Moore, C.L. and Greenblatt, J. (2003) Organization and function of APT, a subcomplex of the yeast cleavage and polyadenylation

factor involved in the formation of mRNA and small nucleolar RNA 3'-ends. *J Biol Chem*, **278**, 33000-33010.

Neuman-Silberberg, F.S., Bhattacharya, S. and Broach, J.R. (1995) Nutrient availability and the RAS/cyclic AMP pathway both induce expression of ribosomal protein genes in Saccharomyces cerevisiae but by different mechanisms. *Mol Cell Biol*, **15**, 3187-3196.

Ng, H.H., Dole, S. and Struhl, K. (2003a) The Rtf1 component of the Paf1 transcriptional elongation complex is required for ubiquitination of histone H2B. *J Biol Chem*, **278**, 33625-33628.

Ng, H.H., Robert, F., Young, R.A. and Struhl, K. (2003b) Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell*, **11**, 709-719.

Ni, Z., Schwartz, B.E., Werner, J., Suarez, J.R. and Lis, J.T. (2004) Coordination of transcription, RNA processing, and surveillance by P-TEFb kinase on heat shock genes. *Mol Cell*, **13**, 55-65.

Nigavekar, S.S., Tan, Y.S. and Cannon, J.F. (2002) Glc8 is a glucose-repressible activator of Glc7 protein phosphatase-1. *Arch Biochem Biophys*, **404**, 71-79.

Ohkura, H., Kinoshita, N., Miyatani, S., Toda, T. and Yanagida, M. (1989) The fission yeast dis2+ gene required for chromosome disjoining encodes one of two putative type 1 protein phosphatases. *Cell*, **57**, 997-1007.

Ohnacker, M., Barabino, S.M., Preker, P.J. and Keller, W. (2000) The WD-repeat protein Pfs2p bridges two essential factors within the yeast pre-mRNA 3'-end-processing complex. *EMBO J*, **19**, 37-47.

Orlova, M., Newlands, J., Das, A., Goldfarb, A. and Borukhov, S. (1995) Intrinsic transcript cleavage activity of RNA polymerase. *Proc Natl Acad Sci U S A*, **92**, 4596-4600.

Orozco, I.J., Kim, S.J. and Martinson, H.G. (2002) The poly(A) signal, without the assistance of any downstream element, directs RNA polymerase II to pause in vivo and then to release stochastically from the template. *J Biol Chem*, **277**, 42899-42911.

Orphanides, G. and Reinberg, D. (2002) A unified theory of gene expression. *Cell*, **108**, 439-451.

Orphanides, G., Wu, W.H., Lane, W.S., Hampsey, M. and Reinberg, D. (1999) The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins. *Nature*, **400**, 284-288.

Pappas, D.L., Jr. and Hampsey, M. (2000) Functional interaction between Ssu72 and the Rpb2 subunit of RNA polymerase II in Saccharomyces cerevisiae. *Mol Cell Biol*, **20**, 8343-8351.

Park, N.J., Tsao, D.C. and Martinson, H.G. (2004) The two steps of poly(A)-dependent termination, pausing and release, can be uncoupled by truncation of the RNA polymerase II carboxyl-terminal repeat domain. *Mol Cell Biol*, **24**, 4092-4103.

Patel, D. and Butler, J.S. (1992) Conditional defect in mRNA 3' end processing caused by a mutation in the gene for poly(A) polymerase. *Mol Cell Biol*, **12**, 3297-3304.

Peggie, M.W., MacKelvie, S.H., Bloecher, A., Knatko, E.V., Tatchell, K. and Stark, M.J. (2002) Essential functions of Sds22p in chromosome stability and nuclear localization of PP1. *J Cell Sci*, **115**, 195-206.

Peng, J., Marshall, N.F. and Price, D.H. (1998) Identification of a cyclin subunit required for the function of Drosophila P-TEFb. *J Biol Chem*, **273**, 13855-13860.

Peng, W.T., Robinson, M.D., Mnaimneh, S., Krogan, N.J., Cagney, G., Morris, Q., Davierwala, A.P., Grigull, J., Yang, X., Zhang, W., Mitsakakis, N., Ryan, O.W., Datta, N., Jojic, V., Pal, C., Canadien, V., Richards, D., Beattie, B., Wu, L.F., Altschuler, S.J., Roweis, S., Frey, B.J., Emili, A., Greenblatt, J.F. and Hughes, T.R. (2003) A panoramic view of yeast noncoding RNA processing. *Cell*, **113**, 919-933.

Penheiter, K.L., Washburn, T.M., Porter, S.E., Hoffman, M.G. and Jaehning, J.A. (2005) A posttranscriptional role for the yeast Paf1-RNA polymerase II complex is revealed by identification of primary targets. *Mol Cell*, **20**, 213-223.

Petfalski, E., Dandekar, T., Henry, Y. and Tollervey, D. (1998) Processing of the precursors to small nucleolar RNAs and rRNAs requires common components. *Mol Cell Biol*, **18**, 1181-1189.

Phatnani, H.P. and Greenleaf, A.L. (2006) Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev*, **20**, 2922-2936.

Pina, B., Fernandez-Larrea, J., Garcia-Reyero, N. and Idrissi, F.Z. (2003) The different (sur)faces of Rap1p. *Mol Genet Genomics*, **268**, 791-798.

Ping, Y.H. and Rana, T.M. (2001) DSIF and NELF interact with RNA polymerase II elongation complex and HIV-1 Tat stimulates P-TEFb-mediated phosphorylation of RNA polymerase II and DSIF during transcription elongation. *J Biol Chem*, **276**, 12951-12958.

Powell, W. and Reines, D. (1996) Mutations in the second largest subunit of RNA polymerase II cause 6- azauracil sensitivity in yeast and increased transcriptional arrest in vitro. *J Biol Chem*, **271**, 6866-6873.

Powers, T. and Walter, P. (1999) Regulation of ribosome biogenesis by the rapamycin-sensitive TOR-signaling pathway in Saccharomyces cerevisiae. *Mol Biol Cell*, **10**, 987-1000.

Preker, P.J., Lingner, J., Minvielle-Sebastia, L. and Keller, W. (1995) The *FIP1* gene encodes a component of a yeast pre-mRNA polyadenylation factor that directly interacts with poly(A) polymerase. *Cell*, **81**, 379-389.

Preker, P.J., Ohnacker, M., Minvielle-Sebastia, L. and Keller, W. (1997) A multisubunit 3'-end processing factor from yeast containing poly(A) polymerase and homologues of the subunits of mammalian cleavage and polyadenylation specificity factor. *The EMBO Journal*, **16**, 4727 - 4737.

Prelich, G. (2002) RNA polymerase II carboxy-terminal domain kinases: emerging clues to their function. *Eukaryot Cell*, **1**, 153-162.

Proudfoot, N. (2000a) Connecting transcription to messenger RNA processing. *Trends Biochem Sci*, **25**, 290-293.

Proudfoot, N. (2000b) Connecting transcription to messenger RNA processing. *Trends Biochem Sci*, **25**, 290-293.

Proudfoot, N. (2004) New perspectives on connecting messenger RNA 3' end formation to transcription. *Curr Opin Cell Biol*, **16**, 272-278.

Proudfoot, N.J. (1989) How RNA polymerase II terminates transcription in higher eukaryotes. *Trends Biochem. Sci.*, **14**, 105-110.

Proudfoot, N.J., Furger, A. and Dye, M.J. (2002) Integrating mRNA processing with transcription. *Cell*, **108**, 501-512.

Qu, L.H., Henry, Y., Nicoloso, M., Michot, B., Azum, M.C., Renalier, M.H., Caizergues-Ferrer, M. and Bachellerie, J.P. (1995) U24, a novel intron-encoded small nucleolar RNA with two 12 nt long, phylogenetically conserved complementarities to 28S rRNA. *Nucleic Acids Res*, **23**, 2669-2676.

Rasmussen, T.P. and Culbertson, M.R. (1998) The putative nucleic acid helicase Sen1p is required for formation and stability of termini and for maximal rates of synthesis and levels of accumulation of small nucleolar RNAs in Saccharomyces cerevisiae. *Mol Cell Biol*, **18**, 6885-6896.

Renner, D.B., Yamaguchi, Y., Wada, T., Handa, H. and Price, D.H. (2001) A highly purified RNA polymerase II elongation control system. *J Biol Chem*, **276**, 42601-42609.

Robinson, L.C., Gibbs, J.B., Marshall, M.S., Sigal, I.S. and Tatchell, K. (1987) CDC25: a component of the RAS-adenylate cyclase pathway in Saccharomyces cerevisiae. *Science*, **235**, 1218-1221.

Rodriguez, C.R., Cho, E.J., Keogh, M.C., Moore, C.L., Greenleaf, A.L. and Buratowski, S. (2000) Kin28, the TFIIH-associated carboxy-terminal domain kinase, facilitates the recruitment of mRNA processing machinery to RNA polymerase II. *Mol Cell Biol*, **20**, 104-112.

Rondon, A.G., Gallardo, M., Garcia-Rubio, M. and Aguilera, A. (2004) Molecular evidence indicating that the yeast PAF complex is required for transcription elongation. *EMBO Rep*, **5**, 47-53.

Rondon, A.G., Garcia-Rubio, M., Gonzalez-Barrera, S. and Aguilera, A. (2003) Molecular evidence for a positive role of Spt4 in transcription elongation. *EMBO J*, **22**, 612-620.

Rudra, D., Zhao, Y. and Warner, J.R. (2005) Central role of Ifh1p-Fhl1p interaction in the synthesis of yeast ribosomal proteins. *EMBO J*, **24**, 533-542.

Russnak, R., Nehrke, K.W. and Platt, T. (1995) *REF2* encodes an RNA-binding protein directly involved in yeast mRNA 3'-end formation. *Molecular and Cellular Biology*, **15**, 1689-1697.

Russo, P. (1995) Saccharomyces cerevisiae mRNA 3' end forming signals are also involved in transcription termination. *Yeast*, **11**, 447-453.

Sachs, A. and Wahle, E. (1993) Poly(A) tail metabolism and function in eucaryotes. *J Biol Chem*, **268**, 22955-25968.

Sadowski, M.D., B.; Hübner, W.; Keller, W. (2003) Independent functions of yeast Pcf11p in pre-mRNA 3' end processing and in transcription termination. *EMBO J,* **22**, 2167-2177.

Sakumoto, N., Mukai, Y., Uchida, K., Kouchi, T., Kuwajima, J., Nakagawa, Y., Sugioka, S., Yamamoto, E., Furuyama, T., Mizubuchi, H., Ohsugi, N., Sakuno, T., Kikuchi, K., Matsuoka, I., Ogawa, N., Kaneko, Y. and Harashima, S. (1999) A series of protein phosphatase gene disruptants in Saccharomyces cerevisiae. *Yeast*, **15**, 1669-1679.

Samarsky, D.A. and Fournier, M.J. (1999) A comprehensive database for the small nucleolar RNAs from Saccharomyces cerevisiae. *Nucleic Acids Res*, **27**, 161-164.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Saunders, A., Core, L.J. and Lis, J.T. (2006) Breaking barriers to transcription elongation. *Nat Rev Mol Cell Biol*, **7**, 557-567.

Schawalder, S.B., Kabani, M., Howald, I., Choudhury, U., Werner, M. and Shore, D. (2004) Growth-regulated recruitment of the essential yeast ribosomal protein gene activator Ifh1. *Nature*, **432**, 1058-1061.

Schmelzle, T., Beck, T., Martin, D.E. and Hall, M.N. (2004) Activation of the RAS/cyclic AMP pathway suppresses a TOR deficiency in yeast. *Mol Cell Biol*, **24**, 338-351.

Shatkin, A.J. and Manley, J.L. (2000) The ends of the affair: capping and polyadenylation. *Nat Struct Biol*, **7**, 838-842.

Shaw, R.J. and Reines, D. (2000) Saccharomyces cerevisiae transcription elongation mutants are defective in PUR5 induction in response to nucleotide depletion. *Mol Cell Biol*, **20**, 7427-7437.

Sheldon, K.E., Mauger, D.M. and Arndt, K.M. (2005) A Requirement for the Saccharomyces cerevisiae Paf1 complex in snoRNA 3' end formation. *Mol Cell*, **20**, 225-236.

Shi, X., Chang, M., Wolf, A.J., Chang, C.H., Frazer-Abel, A.A., Wade, P.A., Burton, Z.F. and Jaehning, J.A. (1997) Cdc73p and Paf1p are found in a novel RNA polymerase II-containing complex distinct from the Srbp-containing holoenzyme. *Mol Cell Biol*, **17**, 1160-1169.

Shi, X., Finkelstein, A., Wolf, A.J., Wade, P.A., Burton, Z.F. and Jaehning, J.A. (1996) Paf1p, an RNA polymerase II-associated factor in Saccharomyces cerevisiae, may have both positive and negative roles in transcription. *Mol Cell Biol*, **16**, 669-676.

Sikorski, R.S. and Hieter, P. (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics*, **122**, 19-27.

Simic, R., Lindstrom, D.L., Tran, H.G., Roinick, K.L., Costa, P.J., Johnson, A.D., Hartzog, G.A. and Arndt, K.M. (2003) Chromatin remodeling protein Chd1 interacts with transcription elongation factors and localizes to transcribed genes. *EMBO J*, **22**, 1846-1856.

Sims, R.J., 3rd, Belotserkovskaya, R. and Reinberg, D. (2004a) Elongation by RNA polymerase II: the short and long of it. *Genes Dev*, **18**, 2437-2468.

Sims, R.J., 3rd, Mandal, S.S. and Reinberg, D. (2004b) Recent highlights of RNA-polymerase-II-mediated transcription. *Curr Opin Cell Biol*, **16**, 263-271.

Sims, R.J., 3rd, Nishioka, K. and Reinberg, D. (2003) Histone lysine methylation: a signature for chromatin function. *Trends Genet*, **19**, 629-639.

Skaar, D.A. and Greenleaf, A.L. (2002) The RNA Polymerase II CTD Kinase CTDK-I Affects Pre-mRNA 3' Cleavage/Polyadenylation through the Processing Component Pti1p. *Mol Cell*, **10**, 1429-1439.

Squazzo, S.L., Costa, P.J., Lindstrom, D.L., Kumer, K.E., Simic, R., Jennings, J.L., Link, A.J., Arndt, K.M. and Hartzog, G.A. (2002) The Paf1 complex physically and functionally associates with transcription elongation factors in vivo. *EMBO J*, **21**, 1764-1774.

Stark, M.J. (1996) Yeast protein serine/threonine phosphatases: multiple roles and diverse regulation. *Yeast*, **12**, 1647-1675.

Steinmetz, E.J. and Brow, D.A. (2003) Ssu72 protein mediates both poly(A)-coupled and poly(A)-independent termination of RNA polymerase II transcription. *Mol Cell Biol*, **23**, 6339-6349.

Steinmetz, E.J., Conrad, N.K., Brow, D.A. and Corden, J.L. (2001) RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts. *Nature*, **413**, 327-331.

Steinmetz, E.J., Ng, S.B., Cloute, J.P. and Brow, D.A. (2006) cis- and trans-Acting determinants of transcription termination by yeast RNA polymerase II. *Mol Cell Biol*, **26**, 2688-2696.

Studitsky, V.M., Walter, W., Kireeva, M., Kashlev, M. and Felsenfeld, G. (2004) Chromatin remodeling by RNA polymerases. *Trends Biochem Sci*, **29**, 127-135.

Sullivan, E.K., Weirich, C.S., Guyon, J.R., Sif, S. and Kingston, R.E. (2001) Transcriptional activation domains of human heat shock factor 1 recruit human SWI/SNF. *Mol Cell Biol*, **21**, 5826-5837.

Swanson, M.S. and Dreyfuss, G. (1988) RNA binding specificity of hnRNP proteins: a subset bind to the 3' end of introns. *EMBO J*, **7**, 3519-3529.

Takagaki, Y. and Manley, J.L. (1997) RNA recognition by the human polyadenylation factor CstF. *Molecular and Cellular Biology*, **17**, 3907-3914.

Tan, S., Aso, T., Conaway, R.C. and Conaway, J.W. (1994) Roles for both the RAP30 and RAP74 subunits of transcription factor IIF in transcription initiation and elongation by RNA polymerase II. *J Biol Chem*, **269**, 25684-25691.

Tanaka, K., Nakafuku, M., Satoh, T., Marshall, M.S., Gibbs, J.B., Matsumoto, K., Kaziro, Y. and Toh-e, A. (1990a) S. cerevisiae genes IRA1 and IRA2 encode proteins that may be functionally equivalent to mammalian ras GTPase activating protein. *Cell*, **60**, 803-807.

Tanaka, K., Nakafuku, M., Tamanoi, F., Kaziro, Y., Matsumoto, K. and Toh-e, A. (1990b) IRA2, a second gene of Saccharomyces cerevisiae that encodes a protein with a domain homologous to mammalian ras GTPase-activating protein. *Mol Cell Biol*, **10**, 4303-4313.

Tarun, S.Z., Jr. and Sachs, A.B. (1996) Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. *EMBO J*, **15**, 7168-7177.

Thevelein, J.M. and de Winde, J.H. (1999) Novel sensing mechanisms and targets for the cAMP-protein kinase A pathway in the yeast Saccharomyces cerevisiae. *Mol Microbiol*, **33**, 904-918.

Thiebaut, M., Kisseleva-Romanova, E., Rougemaille, M., Boulay, J. and Libri, D. (2006) Transcription termination and nuclear degradation of cryptic unstable transcripts: a role for the nrd1-nab3 pathway in genome surveillance. *Mol Cell*, **23**, 853-864.

Thomas, G. and Hall, M.N. (1997) TOR signalling and control of cell growth. *Curr Opin Cell Biol*, **9**, 782-787.

Toda, T., Cameron, S., Sass, P., Zoller, M., Scott, J.D., McMullen, B., Hurwitz, M., Krebs, E.G. and Wigler, M. (1987a) Cloning and characterization of BCY1, a locus encoding a regulatory subunit of the cyclic AMP-dependent protein kinase in Saccharomyces cerevisiae. *Mol Cell Biol*, **7**, 1371-1377.

Toda, T., Cameron, S., Sass, P., Zoller, M. and Wigler, M. (1987b) Three different genes in S. cerevisiae encode the catalytic subunits of the cAMP-dependent protein kinase. *Cell*, **50**, 277-287.

Toda, T., Uno, I., Ishikawa, T., Powers, S., Kataoka, T., Broek, D., Cameron, S., Broach, J., Matsumoto, K. and Wigler, M. (1985) In yeast, RAS proteins are controlling elements of adenylate cyclase. *Cell*, **40**, 27-36.

Tran, H.G., Steger, D.J., Iyer, V.R. and Johnson, A.D. (2000) The chromo domain protein chd1p from budding yeast is an ATP-dependent chromatin-modifying factor. *EMBO J*, **19**, 2323-2331.

Tu, J. and Carlson, M. (1994) The GLC7 type 1 protein phosphatase is required for glucose repression in Saccharomyces cerevisiae. *Mol Cell Biol*, **14**, 6789-6796.

Tu, J. and Carlson, M. (1995) REG1 binds to protein phosphatase type 1 and regulates glucose repression in Saccharomyces cerevisiae. *EMBO J*, **14**, 5939-5946.

Uetz, P., Giot, L., Cagney, G., Mansfield, T.A., Judson, R.S., Knight, J.R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., Qureshi-Emili, A., Li, Y., Godwin, B., Conover, D., Kalbfleisch, T., Vijayadamodar, G., Yang, M., Johnston, M., Fields, S. and Rothberg, J.M. (2000) A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. *Nature*, **403**, 623-627.

Ursic, D., Chinchilla, K., Finkel, J.S. and Culbertson, M.R. (2004) Multiple protein/protein and protein/RNA interactions suggest roles for yeast DNA/RNA helicase Sen1p in transcription, transcription-coupled DNA repair and RNA processing. *Nucleic Acids Res*, **32**, 2441-2452.

Ursic, D., Himmel, K.L., Gurley, K.A., Webb, F. and Culbertson, M.R. (1997) The yeast SEN1 gene is required for the processing of diverse RNA classes. *Nucleic Acids Res*, **25**, 4778-4785.

Valentini, S.R., Weiss, V.H. and Silver, P.A. (1999) Arginine methylation and binding of Hrp1p to the efficiency element for mRNA 3'-end formation. *RNA*, **5**, 1-9.

van Hoof, A., Lennertz, P. and Parker, R. (2000) Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. *Mol Cell Biol*, **20**, 441-452.

Vanacova, S., Wolf, J., Martin, G., Blank, D., Dettwiler, S., Friedlein, A., Langen, H., Keith, G. and Keller, W. (2005) A new yeast poly(A) polymerase complex involved in RNA quality control. *PLoS Biol*, **3**, e189.

Vaquero, A., Loyola, A. and Reinberg, D. (2003) The constantly changing face of chromatin. *Sci Aging Knowledge Environ*, **2003**, RE4.

Vasiljeva, L. and Buratowski, S. (2006) Nrd1 interacts with the nuclear exosome for 3' processing of RNA polymerase II transcripts. *Mol Cell*, **21**, 239-248.

Villa, T., Ceradini, F., Presutti, C. and Bozzoni, I. (1998) Processing of the intron-encoded U18 small nucleolar RNA in the yeast Saccharomyces cerevisiae relies on both exo- and endonucleolytic activities. *Mol Cell Biol*, **18**, 3376-3383.

Wada, T., Takagi, T., Yamaguchi, Y., Ferdous, A., Imai, T., Hirose, S., Sugimoto, S., Yano, K., Hartzog, G.A., Winston, F., Buratowski, S. and Handa, H. (1998a) DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. *Genes Dev*, **12**, 343-356.

Wada, T., Takagi, T., Yamaguchi, Y., Watanabe, D. and Handa, H. (1998b) Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription in vitro. *EMBO J*, **17**, 7395-7403.

Wade, C., Shea, K.A., Jensen, R.V. and McAlear, M.A. (2001) EBP2 is a member of the yeast RRB regulon, a transcriptionally coregulated set of genes that are required for ribosome and rRNA biosynthesis. *Mol Cell Biol*, **21**, 8638-8650.

Wade, J.T., Hall, D.B. and Struhl, K. (2004) The transcription factor Ifh1 is a key regulator of yeast ribosomal protein genes. *Nature*, **432**, 1054-1058.

Wahle, E. and Rüegsegger, U. (1999) 3'-End processing of pre-mRNA in eukaryotes. *FEMS Microbiol Rev*, **23**, 277-295.

Walsh, E.P., Lamont, D.J., Beattie, K.A. and Stark, M.J. (2002) Novel interactions of Saccharomyces cerevisiae type 1 protein phosphatase identified by single-step affinity purification and mass spectrometry. *Biochemistry*, **41**, 2409-2420.

Wang, Y., Pierce, M., Schneper, L., Guldal, C.G., Zhang, X., Tavazoie, S. and Broach, J.R. (2004) Ras and Gpa2 mediate one branch of a redundant glucose signaling pathway in yeast. *PLoS Biol*, **2**, E128.

Warner, J.R. (1999) The economics of ribosome biosynthesis in yeast. *Trends Biochem Sci*, **24**, 437-440.

Weinstein, L.B. and Steitz, J.A. (1999) Guided tours: from precursor snoRNA to functional snoRNP. *Curr Opin Cell Biol*, **11**, 378-384.

Werner-Washburne, M., Braun, E.L., Crawford, M.E. and Peck, V.M. (1996) Stationary phase in Saccharomyces cerevisiae. *Mol Microbiol*, **19**, 1159-1166.

West, S., Gromak, N. and Proudfoot, N.J. (2004) Human 5' --> 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature*, **432**, 522-525.

Wood, A., Schneider, J., Dover, J., Johnston, M. and Shilatifard, A. (2003) The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. *J Biol Chem*, **278**, 34739-34742.

Wu, H., Henras, A., Chanfreau, G. and Feigon, J. (2004) Structural basis for recognition of the AGNN tetraloop RNA fold by the double-stranded RNA-binding domain of Rnt1p RNase III. *Proc Natl Acad Sci U S A*, **101**, 8307-8312.

Wu, X., Hart, H., Cheng, C., Roach, P.J. and Tatchell, K. (2001) Characterization of Gac1p, a regulatory subunit of protein phosphatase type I involved in glycogen accumulation in Saccharomyces cerevisiae. *Mol Genet Genomics*, **265**, 622-635.

Xu, Y.X., Hirose, Y., Zhou, X.Z., Lu, K.P. and Manley, J.L. (2003) Pin1 modulates the structure and function of human RNA polymerase II. *Genes Dev*, **17**, 2765-2776.

Yamada, T., Yamaguchi, Y., Inukai, N., Okamoto, S., Mura, T. and Handa, H. (2006) P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. *Mol Cell*, **21**, 227-237.

Yamaguchi, Y., Wada, T., Watanabe, D., Takagi, T., Hasegawa, J. and Handa, H. (1999) Structure and function of the human transcription elongation factor DSIF. *J Biol Chem*, **274**, 8085-8092.

Yeo, M., Lin, P.S., Dahmus, M.E. and Gill, G.N. (2003) A novel RNA polymerase II C-terminal domain phosphatase that preferentially dephosphorylates serine 5. *J Biol Chem*, **278**, 26078-26085.

Yonaha, M. and Proudfoot, N.J. (1999) Specific transcriptional pausing activates polyadenylation in a coupled in vitro system. *Mol Cell*, **3**, 593-600.

Young, E.T., Kacherovsky, N. and Van Riper, K. (2002) Snf1 protein kinase regulates Adr1 binding to chromatin but not transcription activation. *J Biol Chem*, **277**, 38095-38103.

Zagorski, J., Tollervey, D. and Fournier, M.J. (1988) Characterization of an SNR gene locus in Saccharomyces cerevisiae that specifies both dispensible and essential small nuclear RNAs. *Mol Cell Biol*, **8**, 3282-3290.

Zappacosta, F., Huddleston, M.J., Karcher, R.L., Gelfand, V.I., Carr, S.A. and Annan, R.S. (2002) Improved sensitivity for phosphopeptide mapping using capillary column HPLC and microionspray mass spectrometry: comparative

phosphorylation site mapping from gel-derived proteins. *Anal Chem*, **74**, 3221-3231.

Zawel, L., Kumar, K.P. and Reinberg, D. (1995) Recycling of the general transcription factors during RNA polymerase II transcription. *Genes Dev*, **9**, 1479-1490.

Zhang, Y. (2003) Transcriptional regulation by histone ubiquitination and deubiquitination. *Genes Dev*, **17**, 2733-2740.

Zhang, Z., Fu, J. and Gilmour, D.S. (2005) CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3'-end processing factor, Pcf11. *Genes Dev*, **19**, 1572-1580.

Zhang, Z. and Gilmour, D.S. (2006) Pcf11 is a termination factor in Drosophila that dismantles the elongation complex by bridging the CTD of RNA polymerase II to the nascent transcript. *Mol Cell*, **21**, 65-74.

Zhao, J., Hyman, L. and Moore, C. (1999a) Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol Mol Biol Rev*, **63**, 405-445.

Zhao, Y., McIntosh, K.B., Rudra, D., Schawalder, S., Shore, D. and Warner, J.R. (2006) Fine-structure analysis of ribosomal protein gene transcription. *Mol Cell Biol*, **26**, 4853-4862.

Zhelkovsky, A., Helmling, S. and Moore, C. (1998) Processivity of the Saccharomyces cerevisiae poly(A) polymerase requires interactions at the carboxyl-terminal RNA binding domain. *Mol Cell Biol*, **18**, 5942-5951.

Zorio, D.A. and Bentley, D.L. (2004) The link between mRNA processing and transcription: communication works both ways. *Exp Cell Res*, **296**, 91-97.

10 Acknowledgements

First of all, I would like to thank Prof. Walter Keller, for his enormous generosity, flexibility and support. I am very grateful that he gave me the opportunity to work in his lab and in addition to pursuit a certain “side study”.
Many thanks to all past and present members of the Keller Lab for fruitful discussions, the pleasant atmosphere and the good times we had outside of the lab.
I would also like to thank my collaborators Bertrand and Dietmar for their motivation and discussions making the work on my Ph.D. project very enjoyable.
I am deeply grateful for the catalyzing discussions I had with Nicole Grieder, for her many helpful suggestions while writing this thesis and her friendship.
Georges Martin also gave many helpful suggestions to improve this thesis – thank you very much!
I’m grateful to Anni, Sabine and Claudio for their friendship and encouragement during all these years at the Biozentrum.
Last but not least, I’m very thankful for everything Fam. Schindelin did for me, especially for their casualness with which they adopted me into their family. Thank you so much for your continuous support, you all mean a lot to me.

11 Curriculum Vitae

Name: Röck

First name: Salome

Date and place of birth: November 23 1972 in Ulm, Germany

Nationality: German

Marital status: single

Education and Academic Training

June 1992:	college degree from the Kolleg der Schulbrüder Gymnasium Illertissen, Germany
Oct. 92-July 95:	diploma degree in music (major subject harp), University of Würzburg
Oct. 95-Feb 97:	undergraduate studies in Geology at the University of Basel, Switzerland
Apr. 97-Apr.2001:	studies in Biology at the University of Basel, Switzerland
Sep. 99- Sep 2000:	diploma thesis in the laboratory of Prof. Dr. Markus Affolter, Department of Cell biology, Biozentrum Basel: Functional analysis of *Drosophila* FGF signaling in cell migration
Aug. 2001-Apr. 2002:	internship in the laboratory of Prof. Dr. Walter Keller, Department of Cell biology, Biozentrum, University of Basel, Switzerland
Since Apr. 2002:	Ph.D. student in the laboratory of Prof. Dr. Walter Keller, Department of Cell biology, Biozentrum, University of Basel, Switzerland

Publications

Röck, S., Paguet, B., Dichtl, B., Keller, W. (2007) The function of Glc7p in transcription termination of snoRNAs and elongation of pre-mRNAs. Manuscript in preparation

Röck, S., Dossenbach, C. and Affolter, M. (2001) Specificity of FGF signaling in cell migration in Drosophila. Development 128, 4563-4572