Structure-function analysis of the RNA polymerase III subcomplex C17/25 and genome-wide distribution of RNA polymerase II

Anna Justyna Jasiak

aus Kedzierzyn-Kozle, Polen

2008

Erklärung

Diese Dissertation wurde im Sinne von §13 Abs. 3 der Promotionsordnung vom 29. Januar 1998 von Herrn Prof. Dr. Patrick Cramer betreut.

Ehrenwörtliche Versicherung

Diese Dissertation wurde selbständig und ohne unerlaubte Hilfe erarbeitet.

München, am 21. November 2008

Anna J. Jasiak

Dissertation eingereicht am 21. November 2008

1. Gutachter: Prof. Dr. Patrick Cramer

2. Gutachter: Prof. Dr. Klaus Förstemann

Mündliche Prüfung am 14. Januar 2009

Acknowledgements

I would like to thank my supervisor Prof. Patrick Cramer for creating a highly motivating scientific environment and his open-minded attitude in trying new methods. I have started my PhD with a purely crystallographic project and thanks to his never-ending enthusiasms and ideas I have got a unique opportunity to gain an insight into a broad range of *in vivo* and *in vitro* techniques. I have enjoyed it very much. I am deeply grateful for his support and understanding shown during the last phase of my PhD.

Special thanks to Karim Armache and Laurent Larivière for their help in solving C17/25 structure, to Michaela Bertero for teaching me how to handle the protein crystals and to Dietmar Martin for the discussion and help on the ChIP-chip project.

Moreover, I would like to thank to my collaborators Johannes Söding and Holger Hartmann for their input in microarray data analysis and Birgit Märtens for yeast complementation studies. I wish to express my gratitude to Marian Kalocsay from Prof. Stefan Jentsch group, Eleni Karakasili from Dr. Katja Sträßer laboratory and Angelika Mitterweger from Prof. Peter Becker group for introducing me to the ChIP-chip method.

I am thankful to all present and former members of the Cramer lab for their help and nice atmosphere in the lab. Many thanks Claudia Buchen, Kristin Leike and Stefan Benkert for making the everyday life in the lab easier. I whish to thank to Alessandro Vannini and whole Pol III team for help and inspiring scientific discussions by a self-made tiramisu. Nicole, thank you for being the most wonderful student I could imagine.

I am particularly grateful to Florian Brückner for his help in editing of this thesis.

Dear Erika, Kristin, Gerke, Flo, Bärbel, Vika, Ania, Aga, Nicki, Timo, Isa, Adam, Oliver and Halina, thank you for your friendship and support during the hard times. Without you many things would not be possible.

René - thank you for your patience and support during my writing.

I would like to thank to Alexander, Marita, Wilhelm and Viktor Sternemann for being my German family for almost 4,5 years. Alex - thank you for bringing music into my life.

Asia – thank you… not only for your input in back-upping of this thesis.

Szczególne wyrazy wdzięczności pragnę przekazać mojej Rodzinie. Kochani, dziękuję za waszą miłość, wsparcie i nigdy nieustającą wiarę w moje możliwości.

I wish to dedicate this work to Irena and Jacek Hetper,
my first and most beloved chemistry teachers.

Summary

RNA synthesis in the eukaryote nucleus is carried out by the multisubunit RNA polymerases (Pol) I, II, and III, which comprise 14, 12, and 17 subunits, respectively. All the RNA polymerases share a common architecture, with ten subunits forming a structurally conserved core, and additional subunits located on the periphery of the enzyme. Rpb4/7 complex located on the periphery of RNA Pol II is involved in transcription initiation recognizing of the promoter-bound transcription factors. C17/25 has been suggested to be a functional counterpart of the Rpb4/7 subcomplex in Pol III system. This thesis focuses on structure-function analysis of C17/25 complex in RNA Polymerase III and genome-wide distribution of the Pol II Rpb4/7 subcomplex.

The results presented here provide first structural information on Pol III, the largest nuclear RNA polymerase. We obtained an 11-subunit model of RNA polymerase (Pol) III by combining a homology model of the nine-subunit core enzyme with a new X-ray structure of the subcomplex C17/25. Compared to Pol II, Pol III shows a conserved active center for RNA synthesis, but a structurally different upstream face for specific initiation complex assembly during promoter selection. The Pol III upstream face includes a HRDC domain in subunit C17 that is translated by 35 Å and rotated by 150° compared to its Pol II counterpart. The HRDC domain is essential *in vivo*, folds independently *in vitro*, and, unlike other HRDC domains, shows no indication of nucleic acid binding. Thus the HRDC domain is a functional module that could account for the role of C17 in Pol III promoter-specific initiation. During elongation, C17/25 may bind Pol III transcripts emerging from the adjacent exit pore, since the subcomplex binds to tRNA *in vitro*. These data provide structural insights into Pol III and reveal specific features of the enzyme that can account for functional differences between nuclear RNA polymerases.

Yeast RNA polymerase (Pol) II consists of a ten-subunit core enzyme and the Rpb4/7 subcomplex, which is dispensable for catalytic activity and dissociates *in vitro*. To investigate whether Rpb4/7 is an integral part of DNA-associated Pol II *in vivo*, we used chromatin immunoprecipitation coupled to high-resolution tiling microarray analysis. We show that the genome-wide occupancy profiles for Rpb7 and the core subunit Rpb3 are essentially identical. Thus, the complete Pol II associates with DNA *in vivo*, consistent with functional roles of Rpb4/7 throughout the transcription cycle.

Publications

Parts of this work have been published:

Jasiak, A.J., Armache, K.-J., Martens, B., Jansen, R.-P., Cramer, P., (2006). Structural biology of RNA polymerase III: new Pol III subcomplex X-ray stucture and 11-subunit enzyme model. Mol Cell 23, 71-81.

Cramer, P., Armache, K.-J., aumli, S., Benkert, S., Brueckner, F., Buchen, C., Damsma, G.E., Dengl, S., B Geiger, S.R., Jasiak, A.J., Jawhari, A., Jennebach, S., Kamenski, T., Kettenberger, H., Kuhn, C.-D., Lehmann, E., Leike, K., Sydow, J., Vannini, A. (2008). Structure of Eukaryotic RNA Polymerases. Annu. Rev. Biophys. 37, 337-352.

Jasiak, A. J., Hartmann, H., Karakasili, E., Kalocsay, M., Flatley,A., Kremmer, E., Sträßer, K., Martin, D.E., Söding, J., Cramer P. (2008) Genome-associated RNA Polymerase II includes the dissociable Rpb4/7 subcomplex. J Biol Chem. 283 (39), 26423-7.

Table of contents

Chapter I: General Introduction

1. The flow of genetic information

Genetic information required by all cells to live is stored in DNA and organized in complex genomes. Key steps in retrieving this information involve rewriting of DNA into RNA in a process called transcription and further synthesis of polypeptide chains of protein based on RNA templates during translation events (Figure 1). This unidirectional flow of genetic information was postulated as the central dogma of molecular biology in all organisms (Crick, 1970; Thieffry & Sarkar, 1998). The exception or special case with an inverted flow of the genetic information was found in retroviruses. The RNA-dependent DNA polymerase transfers the information from a viral RNA-based genome to DNA in the process of reverse transcription (Baltimore, 1970; Temin & Mizutani, 1970).

Figure 1: The central dogma of molecular biology

2. DNA-dependent RNA polymerases

Transcription, the first step in decoding of the genetic information is carried out by the multisubunit protein complexes of DNA dependent RNA polymerases. In prokaryotic cells thousands of different genes are transcribed by a common multiprotein machinery. Crystal structures of bacterial RNA polymerases from *Thermus aquaticus* and *Thermus thermophilus* (Vassylyev *et al.*, 2002),(Zhang *et al.*, 1999), as well as an archaeal RNA polymerase (*Sulfolobus solfataricus*) (Hirata *et al.*, 2008) are known. The overall structure consists of a core enzyme including five subunits: β, β', two α and ω. The additional σ subunit was shown to be essential for promoter DNA binding specificity of the polymerase. The two largest subunits β and β' form a crab claw-like structure harbouring the DNA binding channel with the catalytic center containing two Mg^{2+} ions (one of them permanently bound). The two α-subunits are associated one with β and the other with β'-subunit. The small ω-subunit localizes around the C-terminus of β'.

In the much more complex eukaryotic genome the task of transcribing genes is divided among three highly related enzymes, RNA polymerase I, II and III. A fourth type of RNA polymarase was recently discovered in plants (Herr *et al.*, 2005),(Kanno *et al.*, 2005), where it is claimed to play a role in the silencing of repetitive DNA sequences by RNA-directed DNA methylation (Till & Ladurner, 2007).

Each of the eukaryotic RNA polymerases is dedicated to transcription of a specific set of genes and is subject to specific regulation. Pol I is specialized in high-level transcription of rDNA. Thereby it creates a single precursor RNA transcript, which is processed into mature 28S, 5.8S and 18S rRNA. The Pol II transcriptome is much more complex, as it includes not only all different protein-coding mRNAs but also non-coding RNAs such as small nuclear (sn), small nucleolar (sno) or micro (mi) RNAs. The most diverse group of all RNA polymerases transcripts are RNA Pol III products. A wide collection of genes transcribed by this enzyme encodes structural or

Table 1: Subunit composition of multisubunit RNA polymerases

RNA polymerase	Pol I	Pol II	Pol III	Archaea	Bacteria
Ten-subunit core	A190	Rpb1	C160	A' + A"	β'
	A135	Rpb2	C128	B (B' + B")	β
	AC40	Rpb3	AC40	D	α
	AC19	Rpb11	AC19	L	α
	Rpb6 (ABC23)	Rpb6	Rpb6	K	ω
	Rpb5 (ABC27)	Rpb5	Rpb5	H	-
	Rpb8 (ABC14.5)	Rpb8	Rpb8	-	-
	Rpb10 (ABC10β)	Rpb10	Rpb10	N	-
	Rpb12 (ABC10α)	Rpb12	Rpb12	P	-
	A12.2	Rpb9	C11	-	-
Rpb4/7	A14	Rpb4	C17	F	-
complexes	A43	Rpb7	C25	E'	-
TFIIF-like subcomplex[a]	A49	(Tfg1/Rap74)	C37	-	-
	A34.5	(Tfg2/Rap30)	C53	-	-
Pol III-specific subcomplex	-	-	C82	-	-
	-	-	C34	-	-
	-	-	C31	-	-
Number of subunits	14	12	17	11 (12)	5
Molecular weight in kDa (*species*)	589 (*S. cerevisiae*)	514 (*S. cerevisiae*)	693 (*S. cerevisiae*)	380 (*P. furiosus*)	375 (*T. aquaticus*)

[a]The two subunits in Pol I and Pol III are predicted to form heterodimers that resemble part of the Pol II initiation/elongation factor TFIIF, which is composed of subunits Tfg1, Tfg2, and Tfg3 in *Saccharomyces cerevisiae*, and of subunits Rap74 and Rap30 in human.

catalytic RNAs, which are shorter than 400 base pairs. Pol III is dedicated to the synthesis of small RNAs, including transfer RNAs, 5S ribosomal RNA, U6 small nuclear RNA and many other RNAs of known and unknown functions.

Pol I, II, and III comprise 14, 12, and 17 subunits, respectively, and have a total molecular weight of 589, 514, and 693 kDa, respectively. Ten subunits form a structurally conserved core, and additional subunits are located on the periphery. The two largest subunits of the core are homologous to the β and β' subunits of the bacterial RNA polymerases. Five subunits are identical in all three eukaryotic polymerases and two other are shared between Pol I and Pol III systems. All Pol II subunits have counterparts in the other two RNA polymerases, which additionally contain some unique subcomplexes (Table 1). Structural studies have so far concentrated on Pol II. X-ray structures are known of the 10-subunit Pol II core, and of the complete 12-subunit Pol II and several functional complexes (most recently reviewed in (Cramer *et al.*, 2008)). Structural information on the other nuclear RNA polymerases is limited to electron microscopic investigations.

The structural differences between eukaryotic polymerases most probably underlie early observed difference in sensitivity to α-amanitin (Weinmann & Roeder, 1974). α-amanitin is a cyclic octapeptide that can inhibit transcription initiation and elongation by directly binding to the RNA polymerase. Pol II exhibits a high sensitivity to α-amanitin, whereas Pol I activity is not influenced even by high concentrations of the toxin. Pol III shows intermediate sensitivity, but the inhibition occurs at 1000-fold higher α-amanitin concentration than for Pol II. Pol III can be selectively inhibited by a bacterial phytotoxin called tagetitoxin (Steinberg *et al.*, 1990). Detailed structures of core RNA Pol II bound with α-amanitin (Bushnell *et al.*, 2002; Kaplan *et al.*, 2008) and the Pol II elongation complex (Brueckner & Cramer, 2008) unveil the mechanism of α-amanitin inhibition. Polymerase-specific inhibitors are very useful in distinguishing between different polymerase transcripts. Especially facing a steadily growing number of known short non-coding RNAs, a tool allowing discriminating whether they originate form Pol II or other polymerases transcription proofs to be valuable (Nakaya *et al.*, 2007).

3. Transcription mechanism

Synthesis of RNA is one of the basic, evolutionary conserved enzymatic activities in living cells. RNA extension begins with binding of a nucleoside triphosphate (NTP) substrate to the transcription elongation complex (EC) that is formed by the polymerase, DNA, and RNA. Catalytic addition of the nucleotide to the growing RNA 3' end then releases a pyrophosphate ion. Finally, translocation of DNA and RNA frees the substrate site for binding of the next NTP. The EC is characterized by an unwound DNA region, the transcription bubble. The bubble contains a short hybrid duplex formed between the DNA template strand and the RNA product emerging from the active site.

Catalytic nucleotide incorporation apparently follows a two-metal ion mechanism suggested for all polymerases (Steitz, 1998). The Pol II active site contains a persistently bound metal ion (metal A) and a second, mobile metal ion (metal B) (Cramer *et al.*, 2001). Metal A is held by three invariant aspartate side chains and binds the RNA 3' end (Cramer *et al.*, 2001), whereas metal B binds the NTP triphosphate moiety (Westover *et al.*, 2004). Recent studies of functional complexes of the bacterial RNA polymerase revealed the close conservation of the EC structure (Vassylyev *et al.*, 2007) and provided additional insights into nucleotide incorporation (Vassylyev *et al.*, 2007).

Analysis of the known crystallographic structures of the EC of Pol II (Gnatt *et al.*, 2001; Kettenberger *et al.*, 2003; Wang *et al.*, 2006) and bacterial RNA (Vassylyev *et al.*, 2007) polymerase and available biochemical data (Epshtein *et al.*, 2002; Gong *et al.*, 2004; Bar-Nahum *et al.*, 2005) results in a complete model of the nucleotide addition cycle (NAC), reviewed most recently in (Brueckner *et al.*, 2008). Nucleotides are postulated to be incorporated into the nascent RNA chain in a two step mechanism (Kettenberger *et al.*, 2004; Vassylyev *et al.*, 2007). The NTP would first bind to the post-translocation EC in the inactive state to an open active center conformation characterized by an open trigger loop conformation (Figure 2). Delivery of the NTP to the insertion site involves folding of the trigger loop (Wang *et al.*, 2006; Vassylyev *et al.*, 2007), a mobile part of the active center first observed in free bacterial RNA polymerase (Vassylyev *et al.*, 2002), and in the Pol II-TFIIS complex (Kettenberger *et al.*, 2003). Complete folding of the trigger loop leads to closure of the active center, delivery of the NTP to the insertion site, and formation of the catalytically active polymerase conformation. An alternative model for nucleotide delivery involves binding of the NTP to a putative entry site in the pore, in which the nucleotide base is oriented away from the DNA template, and rotation of the NTP around metal ion B directly into the insertion site (Westover *et al.*, 2004). Catalysis of nucleotide incorporation leads to pyrophosphate formation and its release. This enables the trigger loop to adopt the wedged position (Brueckner & Cramer, 2008), presumably

stabilizing a shift in the bridge helix in the direction of translocation. This is belived to accompany and facilitate movement of the DNA-RNA hybrid from pre- to post-translocation position (translocation step 1). During step 2 of translocation the next template base of the downstream DNA positioned in the pretemplating site is twisted by 90º and reaches its templating position in the active center. It can be accomplished by releasing of the wedged position of the trigger loop and relaxation of the bridge helix, which frees the templating site for the next incoming template base. The proposed mechanism probably preserves EC stability during translocation and decreases the energy barrier between pre- and post-translocation states.
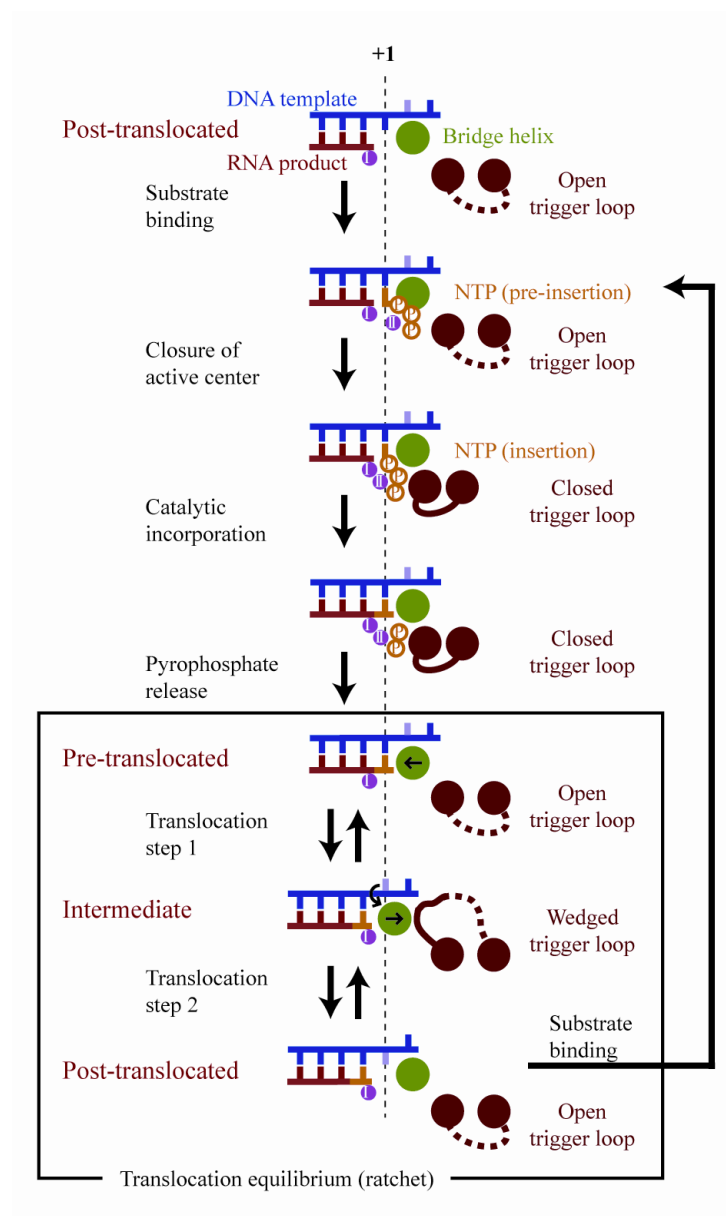
Figure 2. Model of nucleotide addition cycle (NAC). Adapted from (Brueckner *et al.*, 2008).
Schematic representation of the extended model for the NAC. The vertical dashed line indicates register +1. Violet spheres I and II represent metal A and B in the active center, respectively.

Apart from synthesis of the nascent RNA chain, transcription events in the living cell require additional functions of RNA polymerases. During active elongation polymerase has to be able to proofread the transcript recognizing the mismatched nucleotides, as well as overcome the obstacles in DNA template strand. In both cases the enzyme needs to move backward, which results in extrusion of RNA 3' end through the polymerase pore beneath the active site. Backtracking of the DNA-RNA hybrid, causes arrest of the transcription process, which can only be resumed after RNA cleavage and creating a new 3' RNA end. This weak intrinsic endonuclease activity (Wang & Hawley, 1993; Orlova *et al.*, 1995) is stimulated by GreA and GreB proteins in bacterial polymerase (Borukhov *et al.*, 1993) and transcription factor TFIIS in eukaryotic Pol II (Izban & Luse, 1992; Reines, 1992; Kettenberger *et al.*, 2003). In RNA Pol I and Pol III, specific subunits Rpa 12.2 and Rpc11, respectively, seem to take over the role of TFIIS (Chedin *et al.*, 1998; Kuhn *et al.*, 2007).

4. Outline of the thesis

RNA synthesis in the eukaryote nucleus is carried out by the multisubunit RNA polymerases (Pol) I, II, and III, which comprise 14, 12, and 17 subunits, respectively. All the RNA polymerases share a common architecture, with ten subunits forming a structurally conserved core, and additional subunits located on the periphery of the enzyme. Rpb4/7 complex located on the periphery of RNA Pol II has its counterparts in other two RNA Polymerases. C17/25 has been suggested to be a functional counterpart of the Rpb4/7 subcomplex in Pol III system. In all three eukaryotic RNA Polymerases the Rpb4/7-like complexes fulfil important role in transcription initiation recognizing the promoter-bound transcription factors. This thesis focuses on the biology of the Rpb4/7 complexes in the RNA Pol III and Pol II systems. The two separate chapters that correspond to the two papers present the structure-function analysis of C17/25 complex in RNA Polymerase III and genome-wide distribution analysis of the Pol II Rpb4/7 subcomplex, respectively.

The results presented in Chapter II provide first structural information on Pol III, the largest nuclear RNA polymerase. We obtained an 11-subunit model of RNA polymerase (Pol) III by combining a homology model of the nine-subunit core enzyme with a new X-ray structure of the subcomplex C17/25. The functional analysis of the C17/25 complex reveals specific features that can account for functional differences between nuclear RNA polymerases.

In Chapter III the function of Rpb4/7 as an integral part of DNA-associated Pol II *in vivo* is investigated. We show that the genome-wide occupancy profiles for Rpb7 and the core subunit Rpb3, obtained by using chromatin immunoprecipitation coupled to high-resolution tiling microarray analysis, are essentially identical. Thus, the complete Pol II associates with DNA *in vivo*, consistent with functional roles of Rpb4/7 throughout the transcription cycle.

Chapter II: Structure and function of RNA polymerase III C17/25 subcomplex

1. Introduction

1.1. The function of RNA Polymerase III

1.1.1. RNA Polymerase III transcriptome

The central dogma of molecular biology based, on the studies of simple organisms like *Eschericha coli,* has been that RNA functions mainly as an intermediate between a DNA sequence and its encoded protein. The extensive sequences in higher Eukaryotes, which do not encode proteins were for a long time regarded as evolutionary debris accumulated during early assembly of the genes and due to insertions of mobile genetic elements. Evidence accumulating during the last decade of research influenced this view strongly, showing important functions of non-coding (nc) RNAs (Mattick & Makunin, 2006). However until recently the main focus of research on Pol III transcription was on its role in supplying RNAs involved in mRNA translation or small nuclear (sn) RNAs involved in splicing. Indeed, actively dividing cells dedicate about three-quarters of their transcription to produce RNA of the translation machinery. In case of Pol II, almost 50% of its transcription is absorbed by ribosome biosynthesis (Warner, 1999). But as RNA Pol II transcription is not exclusively devoted to ribosomal proteins, the Pol III transcriptome embraces a wide range of non-tRNA and non-rRNA transcripts of known and unknown functions.

The list of genes underlying Pol III transcription was published in the 1990s (White, 1998), (Willis, 1993) and reviewed later including more recent studies (Dieci *et al.*, 2007) (Table 1.1).

Table 1.1 Selected RNA Polymerase III transcripts.

RNA	Species	Reference
tRNA	various species	(Sprinzl *et al.*, 1991)
5S RNA	various species	(Specht *et al.*, 1991)
U6 snRNA	various species	(Willis, 1993)
RNase P RNA	various species	(Willis, 1993)
RNase MRP RNA	*H.sapiens,* rodents	(Willis, 1993)
7SL RNA	various species	(Gupta & Reddy, 1991)
valut RNA	various species	(van Zon *et al.*, 2003)
Y RNA	*H.sapiens*	(Wolin, 1985)
7SK RNA	*H.sapiens*	(Blencowe, 2002)
BC1 BC200 RNA	*H.sapiens,* rodents	(Martignetti & Brosius, 1993; Martignetti & Brosius, 1995)
snaRNAs	*H.sapiens*, chimpanzees	(Parrott & Mathews, 2007)
Telomerase RNA	*Tetrahymena*	(Romero & Blackburn, 1981)
G8 RNA	*Tetrahymena*	(Hallberg *et al.*, 1992)
Myc exon 1	*H.sapiens*	(Sussman *et al.*, 1991)
VA RNA	*Adenovirus*	(Dieci *et al.*, 2007)
EBER RNA	Epstein-Barr virus	(Gupta & Reddy, 1991)
HVP RNA	*Herpesvirus*	(Gupta & Reddy, 1991)
stem-bulge (sb)RNA	*C.elegans*	(Deng *et al.*, 2006)
snoRNAs	*A.thaliana, D.melanogaster, C.elegans, S.cerevisiae*	(Dieci *et al.*, 2007)
miRNAs	*H.sapiens, Herpesvirus*	(Pfeffer *et al.*, 2005; Borchert *et al.*, 2006)
regulatory ncRNAs	*H.sapiens*	(Pagano *et al.*, 2007)
regulatory ncRNAs from TRF1/BRF associated loci	*D.melanogaster*	(Isogai *et al.*, 2007)
Epstein-Barr virus-induced ncRNA	*H.sapiens*	(Mrazek *et al.*, 2007)
SINE RNAs	*H.sapiens,*	(Dieci *et al.*, 2007)

1.1.2. Role of Pol III transcription in the cell

Increasing understanding of the regulatory processes provides a complex picture illustrating the role of the RNA Pol III transcripts in the cell (Figure 1.1) (Dieci *et al.*, 2007). Some Pol III products, like 7SK RNA in vertebrates, are able to regulate transcription by repressing the Pol II elongation, whereas others facilitate post-transcriptional splicing like the U6 snRNA. Pol III transcribes also a multitude of "house-keeping RNAs" involved in the protein biosynthesis. To this group belong various tRNAs and 5S rRNA, which are part of a translating ribosome, but also

RNase P RNA, involved in maturation of pre-tRNAs by processing of its 5' termini, or RMRP RNA and snoRNAs, which play a role in rRNA maturation. Further, BC1 and BC200 RNAs are postulated to be implicated in translation of dendric mRNAs in rodents and primates, respectively, and adenovirus VA-I and VA-II RNAs were shown to inhibit the protein kinase PKR, which is activated in response to viral infection, and allow the translation of the adenoviral mRNAs. The 7SL RNA provides a scaffold for the signal recognition particle and therefore takes part in the post-translational events of protein biosynthesis. Other cytoplasmic RNAs are so-called valut (v)RNAs , which are a part of a large ribonucleoprotein complexes found in many eukaryotes. They were proposed to be involved in multidrugresistance of human tumors and in macromolecular assembly and transport. Pol III transcripts localize also in mitochondria. RNase MRP associated RNA plays a role in the RNA primer processing during mitochondrial DNA replication. A group possessing an enormous regulatory potential is represented by SINE encoded RNAs. Short interspersed repeated DNA elements (SINE) are nonautonomous retrosposons originating from Pol III-transcribed genes, like Alu and diverse micro (mi)RNAs. Their expression is normally low but can be simulated by cell stress, heat shock or viral infection. Mouse B2 RNA can inhibit transcription by direct binding to Pol II, and Alu RNAs can modulate protein translation and downregulate protein expression through an antisense effect.
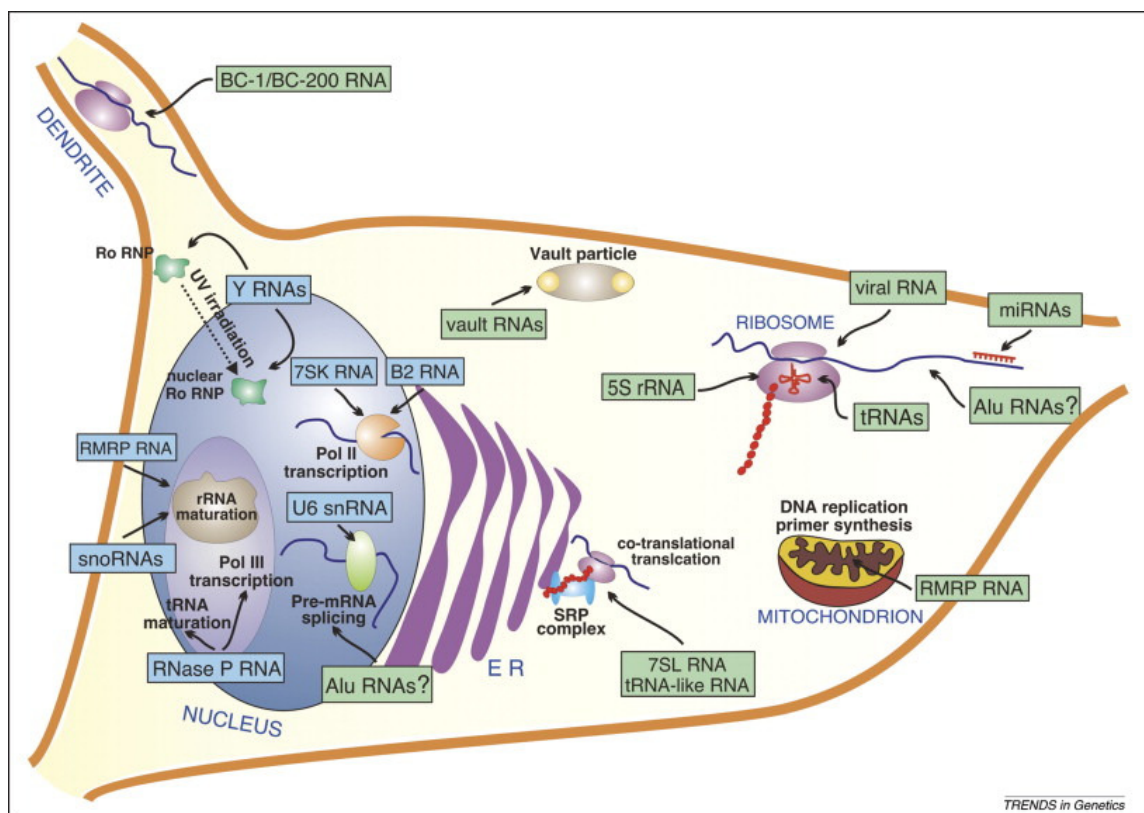
Figure 1.1 Cellular roles of RNA Polymerase III transcripts. Adapted from (Dieci *et al.*, 2007)

1.1.3. Pol III-associated disorders

An increasing number of ncRNAs have been associated with human disorders. Pol III transcripts are involved in diverse cellular processes (Chapter 1.1.2). Therefore alteration in levels or functions of Pol III-synthesised RNA plays a key role in many diseases.

Growth-dependent Pol III transcription of structural RNAs like tRNA and 5S RNA is deregulated in cancer (White, 2004; White, 2005). Regulation of transcription in the cell helps to ensure that the output of Pol III is restricted under conditions that are inappropriate for growth. Transcription can be activated following mitogenic stimulation during entry of the cell cycle (late G(1) phase) (Scott *et al.*, 2001). In the untransformed cells tumour suppressor proteins p53 and Rb directly bind to the TFIIIB, an essential Pol III initiation factor, which in result is unable to interact with TFIIIC or Pol III (White, 2004; White, 2005) (Figure 1.2). Compromised function of these proteins, is it through mutation, hyperphosphorylation or binding to viral oncoproteins, is shown for many human malignancies and is believed to be required for a cancer to develop (Hanahan & Weinberg, 2000). The TFIIIB transcription factor is also a target for oncogenic proteins that stimulate its activity (Figure 1.2). The kinase CK2 and mitogen-activated protein kinase (MAPK) Erk phosphorylate TFIIIB subunits and enhance its ability to bind both Pol III and TFIIIC. In addition Erk kinase phosphorylates and stabilises the oncoprotein c-Myc, significantly increasing its concentration. c-Myc interacts with TFIIIB and has a potent
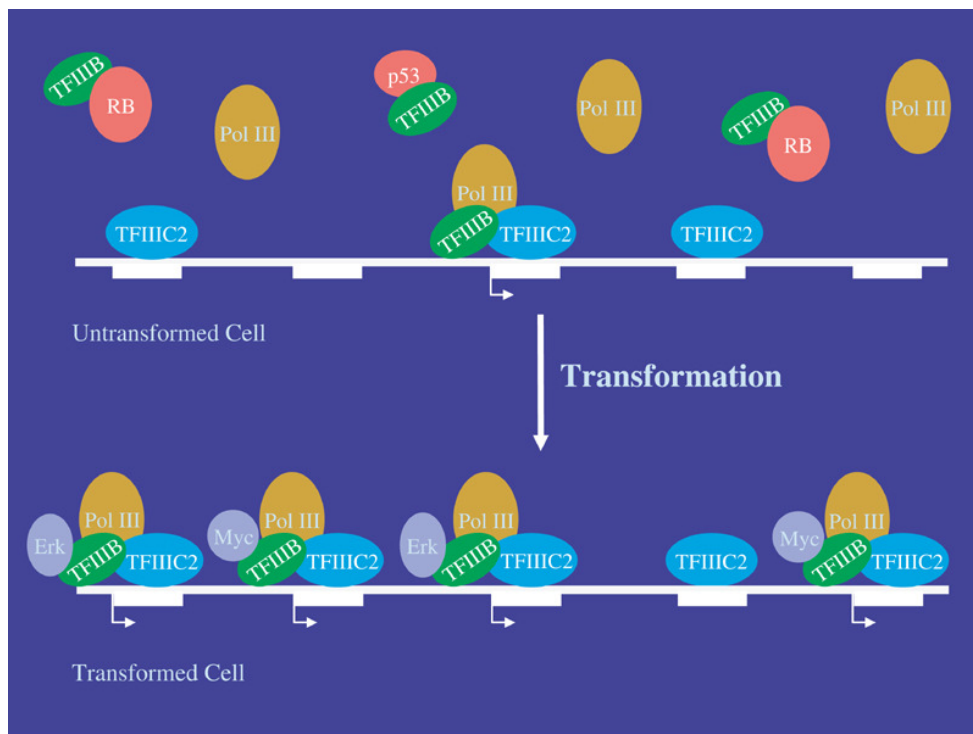
Figure 1.2 Model of Pol III upregulation in the tumor cells. Adapted from (White, 2004).

stimulatory effect on transcription in mouse and human cells (Gomez-Roman *et al.*, 2003). The activation of Pol III transcription can also occur directly, through an increased production of the transcription factors. Transcription factor TFIIIC2 is found to be overexpressed constantly in ovarian carcinomas. Examined tumours displayed elevated TFIIIC2 activity and mRNA levels encoding five subunits of this factor were increased (Winter *et al.*, 2000).

In addition to important role in the initiation and progression of cancer, changes in diverse Pol III genes underlie several human genetic maladies. Mutations in RNase MRP RNA gene region are associated with Cartilage hair hypoplasia (CHH) (Ridanpaa *et al.*, 2001), in which the patients are prone to T cell-associated malignancies, Omenn syndrome, a severe combined immunodeficiency (Roifman *et al.*, 2006), or Schmid type metaphyseal chondrodysplasia (MCDS), a cartilage/bone-related disorder (Ridanpaa *et al.*, 2003). Specific primate small cytoplasmic BC200 RNA and BC1, its functional counterpart in rodents, are postulated to be involved in the regulation of protein translation in well-defined brain regions, such as hippocampus and cortical neurons and associate with Fragile X mental retardation syndrome (Zalfa *et al.*, 2003). Upregulation of BC200 expression seems to play a role in the increased production of amyloid β peptide in Alzheimer disease (Mus *et al.*, 2007).

Future studies of factors specifically regulating Pol III transcription of ncRNAs or their tissue-specific interaction partners as well as disease-related mutations in Pol III transcripts would be of great interest in unveiling the etiology of many human disorders.

1.2. RNA Pol III transcription cycle

1.2.1. Class III promoters

For the diverse types of transcribed RNAs three main types of the RNA polymerase III-specific promoters have been identified (reviewed in (Geiduschek & Kassavetis, 2001; Schramm & Hernandez, 2002; Dieci *et al.*, 2007)). The promoters of most Pol III-transcribed genes include internal control regions (ICRs), which are discontinuous intragenic structures composed of essential sequence blocks separated by nonessential nucleotides. ICR sequences are highly conserved between different genes and different species. One of the first characterised Pol III promoters where those of the *Xenopus laevis* 5S rRNA genes (Bogenhagen *et al.*, 1980; Sakonju *et al.*, 1980). They are referred to as type-1 promoters and their ICR comprise three distinct elements: an A-block, an intermediate element (IE) and a C-block. Most class III genes, including tRNA, VA, Alu, EBER, 7SL, B1, and B2 genes, have type-2

promoters composed of an intragenic A- and B-block. The A-block is located much further from the start site in type I than it is in type II promoters and the spacing between the A- and B-boxes varies greatly, partially due to an intron accommodation. The strength of type I and II promoters can also be affected by extragenic sequences, which in contrast to ICR show frequently little or no conservation. In the type-3 promoters, such as those of the mammalian U6 and 7SK genes, transcription is independent of intragenic elements and is dictated solely by gene external promoters (Murphy *et al.*, 1986; Das *et al.*, 1988; Kunkel & Pederson, 1988). They are located in the 5'-flanking region of the gene and consist of a proximal sequence element (PSE) and a downstream TATA box. The expression of the type-3 promoter genes is enhanced by a distal sequence element (DSE), which can contain several protein binding sites, like an SPH element or an octamer sequence. Additional to the three described types, promoters with both gene external and internal elements were found (Dieci *et al.*, 2000; Ouyang *et al.*, 2000; Yukawa *et al.*, 2000; Giuliodori *et al.*, 2003). An example of such a hybrid promoter in *S. cerevisiae* is an U6 snRNA promoter. It consists of an A box, a B box located at an unusual position 120 bp downstream of the RNA coding region, and a TATA box located upstream of the transcription start site. All three of these promoter elements are required for efficient transcription in vivo (Brow & Guthrie, 1990; Eschenlauer *et al.*, 1993). A schematic illustration of all types of class III promoters is provided in Figure 1.3.
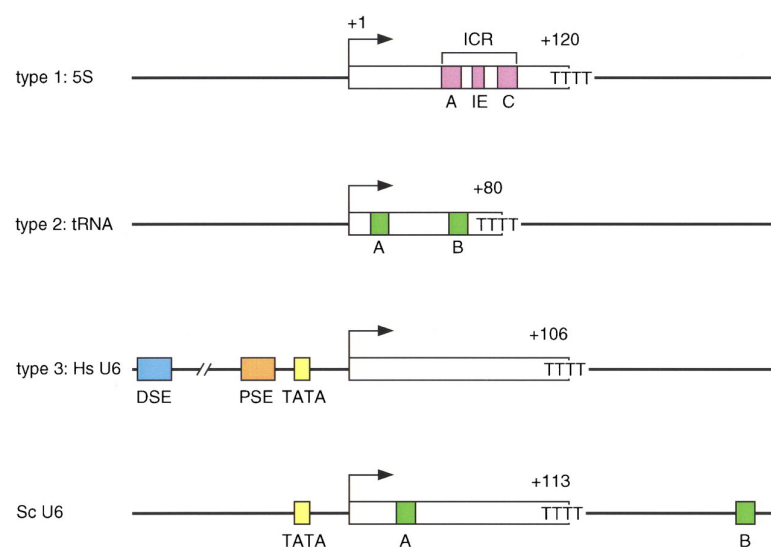
Figure 1.3 RNA Polymerase III promoters. Adapted from (Schramm & Hernandez, 2002).

1.2.2. Initiation complex assembly

In comparision to other eukaryotic polymerases that require numerous transcription factors to form a stable preinititiation complex, the RNA Pol III system is quite simple. The assembly pathway for initiation complex formation depends on the promoter type (Schramm & Hernandez, 2002). In the type-1 promoters ICR is recognized by a zinc-

finger-protein, referred to as TFIIIA (Engelke *et al.*, 1980; Sakonju *et al.*, 1981) (Figure 1.4 b). TFIIIA–DNA complex formation allows for the binding of TFIIIC (Lassar *et al.*, 1983), and the subsequent recruitment of TFIIIB, a transcription factor composed of three subunits, the TATA-box-binding-protein (TBP), Brf1 and Bdp1 (Bartholomew *et al.*, 1991; Kassavetis *et al.*, 1991). TFIIIB binding to the promoter in turn allows the recruitment of RNA polymerase III. In case of the type-2 promoters the A- and B-boxes are recognized directly by TFIIIC (Lassar *et al.*, 1983) (Figure 1.4 a). After the binding of TFIIIC, the pathway to recruitment of the polymerase is similar to that in type 1 promoters, with the subsequent recruitment of TFIIIB and RNA polymerase III. In type-3 promoters, the PSE is recognized by a multisubunit complex called the snRNA activating protein complex (SNAPc) (Figure 1.4 c). The TATA box is recognized by the TBP component of a specialized TFIIIB-like activity (Schramm *et al.*, 2000; Teichmann *et al.*, 2000). The binding of SNAPc and the TFIIIB-like activity then leads to recruitment of RNA polymerase III. Interestingly, a recruitment pathway in which TFIIIB is directly recruited to a TATA box can be observed in vitro with *S. cerevisiae* TFIIIB (Mitchell *et al.*, 1992) (Figure 1.4 d). This reveals a profound aspect of RNA polymerase III transcription, namely, that TFIIIB is sufficient to recruit RNA polymerase and direct several rounds of transcription. Thus, TFIIIA, TFIIIC and SNAPc can be viewed as recruitment factors whose main function is to guide TFIIIB to different promoters, which then allows the recruitment of RNA polymerase III.
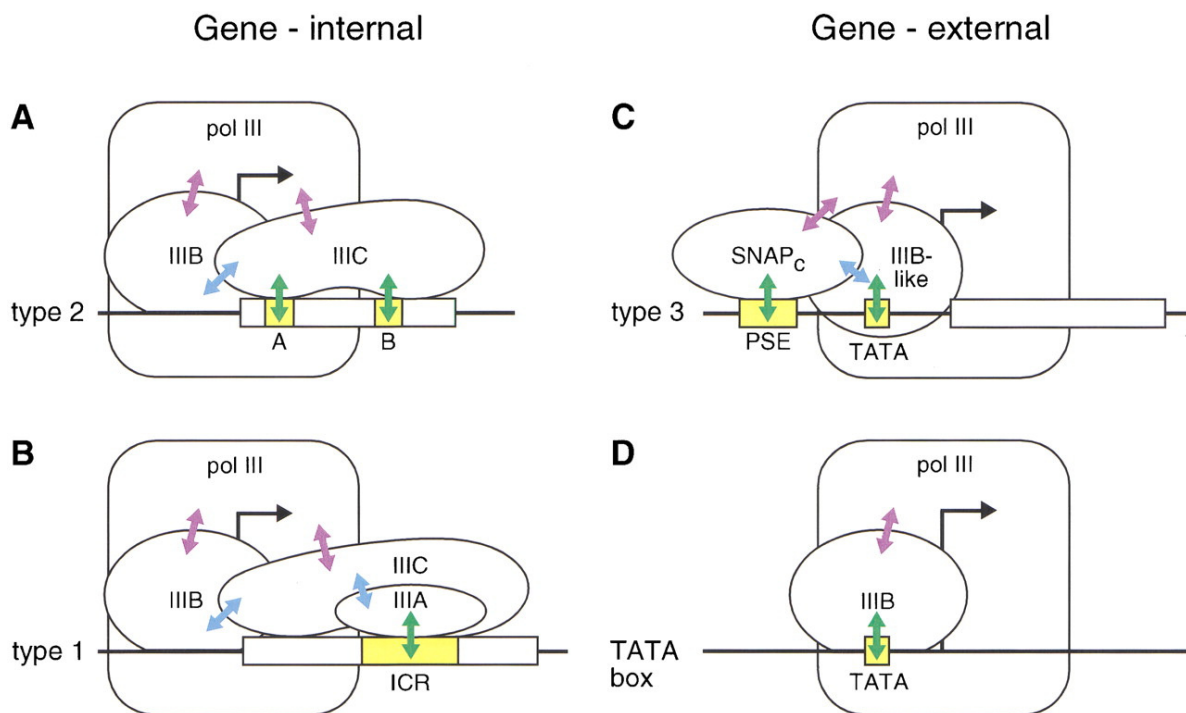
Figure 1.4 Different pathways for recruitment of Pol III to the promoter. Adapted from (Schramm & Hernandez, 2002).

1.2.3. Transcription elongation, termination and reinitiation

During the initial phase of RNA chain elongation, Pol III, just as other RNA polymerases, abortively produces short transcripts (Bhargava & Kassavetis, 1999). After successful incorporation of several ribonucleotides, promoter clearance occurs and Pol III severs its connection to TFIIIB (Kassavetis *et al.*, 1992). Short length of Pol III-transcribed genes results in proximity of transcription start and termination sites. This fact and remarkable stability of TFIIIB-DNA complexes, even in the absence of TFIIIC, are made use of in facilitated transcription reinitiation (Dieci & Sentenac, 1996; Dieci & Sentenac, 2003; Ferrari *et al.*, 2004). RNA Pol III is able to reassemble with the TFIIIB-marked promoter of previously transcribed gene during several rounds of transcription without even releasing the gene after reaching the terminator. Bypassing of the TFIIIC and TFIIIB recruitment steps in preinitiation-complex formation and accommodation of multiple polymerases per gene contributes to the known high transcription rate of RNA Pol III genes (French *et al.*, 2008). RNA chain elongation proceeds unevenly, and Pol III seems to need particularly much time when three UMP residues in succession have to be added (Matsuzaki *et al.*, 1994). Such pausing seems to be crucial for termination, which is accompanied by generation of short oligonucleotides by hydrolytic cleavage at the RNA 3′ end (Bobkova & Hall, 1997). In contrast to other eukaryotic polymerases RNA Pol III is capable of recognizing a termination signal, provided by a simple run of T residues, in a transcription factor-independent manner (Geiduschek & Kassavetis, 2001; Dieci *et al.*, 2007). Recent studies show the critical role of the C53 and C37, as well as C11 subunits of Pol III for transcription termination and reinitiation, recpectively (Landrieux *et al.*, 2006).

1.2.4. Pol III transcription regulation – role of the Maf1 protein

In yeast, transcription activity is regulated according to the growth conditions by the nutrient-sensing signal transduction cascades RAS and TOR. Regulation of the Pol III transcription machinery by diverse signalling pathways is mediated by the Maf1 protein (Upadhya *et al.*, 2002). Maf1 is a hydrophilic phosphoprotein conserved from yeast to human and is considered to be a general suppressor of Pol III transcription (Pluta *et al.*, 2001; Willis & Moir, 2007). In pull-down experiments *S.cerevisiae* Maf1 binds the N-terminal domain of Rpc160 and the Brf1 subunit of TFIIIB, though weakly (Desai *et al.*, 2005; Oficjalska-Pham *et al.*, 2006). Similar results are obtained for the human homologue of this protein (Reina *et al.*, 2006; Rollins *et al.*, 2007). Genetic interactions suggested that Maf1 targets the C31 subunit of Pol III (Ciesla & Boguta, 2008). Based on this suggestion a model of Pol III repression, in which C31-bound Maf1 hampers Pol III interaction with TFIIIB, affecting transcription initiation, was

proposed. Only the dephosphorylated form of Maf1 is shown to bind to Pol III and inhibit transcription (Oficjalska-Pham *et al.*, 2006; Reina *et al.*, 2006; Roberts *et al.*, 2006). The intracellular localization of Maf1 seems to be correlated with its phosphorylation state. Dephosphorylated Maf1 is found in the nucleus, facilitating Pol III repression, whereas inactive, phosphorylated form localizes in the cytoplasm (Oficjalska-Pham *et al.*, 2006; Roberts *et al.*, 2006; Ciesla *et al.*, 2007). Figure 1.5 provides an overview of a current model of transcription regulation by Maf1 (Ciesla & Boguta, 2008). In actively growing cells Maf1 is phosphorylated and localized predominantly in the cytoplasm. Under repressing conditions Maf1 is targeted for dephosphorylation by PP2A phosphatase of the TOR signalling pathway and is imported into the nucleus, where repression of the Pol III activity occurs (Oficjalska-Pham *et al.*, 2006). Favourable growth conditions in the presence of glucose promote derepression of Pol III. Maf1 is phosphorylated by PKA kinase from the RAS pathway (Moir *et al.*, 2006) or some other unknown kinase (Ciesla *et al.*, 2007), dissociate from Pol III and is subsequently exported out of the nucleus by the Man5 carrier (Towpik *et al.*, 2008).
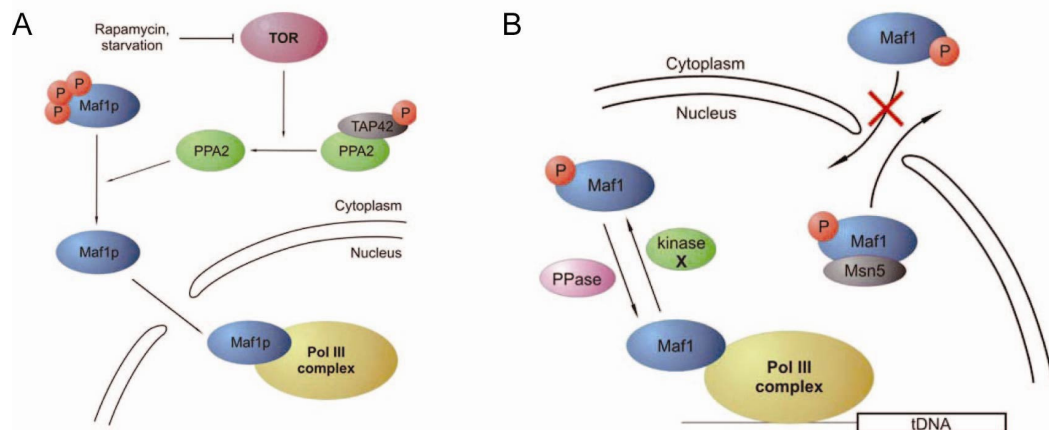
Figure 1.5 Model of Maf1 regulation. Adapted from (Ciesla & Boguta, 2008).

A. Repression of Pol III by Maf1. Maf1 is activeted by dephosphorylation and imported to the nucleus.

B. Derepression of Pol III. Phosphorylated Maf1 is exported out of the nucleus by the Msn5 carrier.

1.3. RNA Polymerase structure

1.3.1. Overall architecture and subunit composition

Pol III is the most complex of all three eukaryotic RNA polymerases. It has a total molecular weight of 693 kDA and compromises 17 subunits. In contrast with the other two nuclear RNA polymerases, all of the Pol III subunits were shown to be essential for yeast viability. Compared with detailed structural studies of Pol II (Cramer *et al.*, 2008) structural information on Pol III is very limited. At the time these

studies were initiated an protein-protein interaction map between Pol III subunits, derived by yeast two-hybrid analysis and copurification assays, as well as some functional analysis concerning single enzyme subunits were available (for reviews see (Chedin *et al.*, 1998; Geiduschek & Kassavetis, 2001; Schramm & Hernandez, 2002). RNA Pol III subunits range from 10 to 160 kDa in size. 10 of them are unique to RNA polymerase III and are designated the C subunits, two, AC40 and AC19, are common to RNA polymerases I and III and therefore designated AC subunits, and five are common to the three RNA polymerases and are designated ABC subunits or referred to as Pol II subunits (Rpb5, Rpb6, Rpb8, Rpb10, and Rpb12). The RNA Pol III overall architecture is characterized by a structural core formed of ten subunits and pheripherical orientated subcomplexes (see Chapter I, Table 1). The subunit composition is conserved from yeast to human and originated on a very early stage of the evolution (Huang & Maraia, 2001; Hu *et al.*, 2002; Proshkina *et al.*, 2006). The two largest subunits, C160 and C128, show substantial homology to the Pol II subunits Rpb1 and Rpb2, and E.coli RNA polymerase β and β' subunits, respectively. AC40 and AC19 are homologous to the Pol II subunits Rpb3 and Rpb 11, and α_2 of bacterial RNA Polymerase. ABC23 (Rpb6) is evolutionarily related with the ω core subunit of the bacterial enzyme. C11 shows limited homology to the Pol II subunit Rpb9 and, in a C-terminal segment, to the Pol II elongation factor TFIIS (Chedin *et al.*, 1998; Kettenberger *et al.*, 2003). It has been also shown to be implicated in nascent RNA hydrolysis coupled to polymerase backtracking along its DNA template (Landrieux *et al.*, 2006). On the periphery of the core enzyme, Pol III contains seven additional subunits, which form three distinct subcomplexes. The subcomplex C82/34/31 (Werner *et al.*, 1992; Wang & Roeder, 1997) and the subcomplex C53/37 (Hu *et al.*, 2002; Landrieux *et al.*, 2006) are Pol III-specific. The third subcomplex, C17/25 has been suggested to be the counterpart of subcomplexes Rpb4/7 in Pol II (Sadhale & Woychik, 1994; Hu *et al.*, 2002; Siaut *et al.*, 2003), Rpa14/43 in Pol I (Shpakovski & Shematorova, 1999; Peyroche *et al.*, 2002; Meka *et al.*, 2003), and RpoF/E in archaeal RNA polymerase (Todone *et al.*, 2001), although the corresponding subunit sequences show only weak conservation in some regions. Recent mass spectrometry studies on *S. cerevisiae* Pol III support pheripherical location of these three subcomplexes (Lorenzen *et al.*, 2007).

1.3.2. RNA Pol III subcomplexes and its role in transcription

1.3.2.1. C17/25

C25 is a highly conserved subunit of RNA Pol III with homology to Rpa43 (Pol I), Rpb7 (Pol II) and RpoE (archeal RNA polymerase) (Sadhale & Woychik, 1994; Shpakovski & Shematorova, 1999). C17 was one of the last subunits to be identified as part of the Pol III multisubunit complex (Ferri *et al.*, 2000). In contrast to C25, identification of C17 as an Rpb4 counterpart in the RNA Pol II system came much later (Siaut *et al.*, 2003). Although C17 is well conserved between different species it shares very weak sequence conservation with its counterparts in other polymerases (Siaut *et al.*, 2003). In RNA Pol I and the archeal RNA polymerase, C17 is a counterpart of Rpa14 and RpoF, respectively (Peyroche *et al.*, 2002; Meka *et al.*, 2003). The Pol III subunits C25 and C17 form a heterodimeric subcomplex similar to the Pol II paralogs Rpb7 and Rpb4. A model of the three-dimensional structure of C17/25 was constructed based on the sequence alignments and using the RpoE/F X-ray structure (Siaut *et al.*, 2003).

Like all Pol III subunits, both C17 and C25 are essential for viability in yeast (Sadhale & Woychik, 1994; Ferri *et al.*, 2000). Similar to Rpb4/7, the C17/25 subcomplex is involved in transcription initiation recognizing of the promoter-bound factors, but is not required for transcription `elongation or termination (Zaros & Thuriaux, 2005). Yeast two-hybrid screen and coimmuno-precipitation experiments indicate that C17 interacts with the N-terminal part of Brf1, a subunit of the transcription initiation factor TFIIIB. C17 interacts also with the RNA polymerase III C31 subunit (Ferri *et al.*, 2000), which is itself required for transcription initiation (Werner *et al.*, 1992; Werner *et al.*, 1993; Wang & Roeder, 1997).

1.3.2.2. C31/34/82

C82, C34 and C31 form a stable subcomplex in yeast and human (Werner *et al.*, 1992; Wang & Roeder, 1997). The C31/34/82 subcomplex is detachable from the enzyme bearing mutation in the N-terminal zinc-binding domain of the largest subunit C160 (Werner *et al.*, 1992). Recent mass spectrometry analysis of the RNA Pol III architecture revealed that the C31 subunit bridges between the C82/34 dimer, the Pol III core and the C17/25 subcomplex (Lorenzen *et al.*, 2007).

C31/34/82 seems to be involved in specific promoter recognition and therefore is required for a specific transcription initiation (Thuillier *et al.*, 1995). C34 mapping revealed that its position is most upstream on the promoter of all the Pol III subunits (Bartholomew *et al.*, 1993). It was proposed to have a dual role in Pol III recruitment and in open complex formation (Brun *et al.*, 1997). Consistent with this observation,

the C34 subunit interacts directly with the Brf1 subunit of TFIIIB, the key transcription factor that recruits Pol III to specific promoters (Werner *et al.*, 1993; Khoo *et al.*, 1994; Andrau *et al.*, 1999). Similarly, the human homolog of C34, hRPC39(RPC6), was found to physically interact with human TATA-binding protein and hBrf1, both subunits of human TFIIIB (Wang & Roeder, 1997). C31 interacts with C17, another Pol III subunit that also binds to Brf1 (Ferri *et al.*, 2000). This two fold strengthened interaction of Pol III subcomplexes with the Pol III-specific transcription factors could help loading the whole enzyme on its large pre-initiation complex (Geiduschek & Kassavetis, 2001).

1.3.2.3. C37/53

Subunits C53 and C37 form a heterodimer stably associated with the Pol III core (Lorenzen *et al.*, 2007), but are present in subtechiometric amounts. C37 and C53 are homologues to the Pol I Rpa49 and Rpa34.5 subunits, respectively (Kuhn *et al.*, 2007; Cramer *et al.*, 2008). Structural similarity prediction (HHpred) showed also a slight homology to TFIIF-α and TFIIF-β, respectively. Based on the observation that mutations in either C37 or C11 lead to the loss of C53, C37 and C11 after Pol III purification, the C53/37 subcomplex was proposed to form an autonomous structural module with C11 (Hu *et al.*, 2002; Landrieux *et al.*, 2006). The fact that the association of these three subunits in isolation was never demonstrated together with the mass spectrometry data studying the architecture of the RNA Pol III (Lorenzen *et al.*, 2007) suggest that C11 belongs to the enzyme core, even though a trimeric subcomplex C53/37/11 cannot be completely ruled out.

Experiments with an RNA polymerase III mutant, lacking subunits C53, C37 and C11 (PolIIIΔ) enlightened their specific functions during transcription (Landrieux *et al.*, 2006). The C53 and C37 subunits are not required for the basal process of RNA synthesis but seem to be crucial for the efficient termination of the transcription. Moreover, the addition of the C53/C37 complex reduced the elongation rate of PolIIIΔ to the rate observed with wildtype RNA polymerase III. It supports the suggestion that the natural pause at the terminator is important for termination (Gusarov & Nudler, 1999; Bar-Nahum *et al.*, 2005). The same study showes that correct terminator recognition is not sufficient to allow Pol III transcription re-initiation. This important function is directed solely by the C11 subunit of Pol III (Landrieux *et al.*, 2006).

1.4. Aim of this study

High-resolution structural information on eukaryotic RNA polymerases is limited to Pol II. Being the largest of all three polymerases, RNA pol III poses a challenge for structural analysis. Until now no crystals of the whole enzyme could be obtained. As observed before by studying RNA Pol II, the Mediator complex or RNA Pol I, gaining insights into the architecture of macromolecular complexes proofs to be a time consuming and a complicated task. For those large complexes, atomic structures of single subunits or subcomplexes were determined and combined with the lower resolution EM or crystallographic data of the whole multisubunit complex (Armache *et al.*, 2005; Kuhn *et al.*, 2007). The aim of this work was to gain first insights into the yeast RNA Polymerase III structure and provide a framework for comparative structural and functional analysis of eukaryotic polymerases. This task was achieved by crystallization of the Pol III C17/25 subcomplex combined with functional analysis *in vivo* and *in vitro*. C17/25 is a paralog of Rpb4/7 in the RNA Pol II system, even though it shows a relatively weak amino acid sequence identity. The functional information on this subcomplex was limited to its role in promoter-dependent transcription initiation.

2. Results

2.1. Expression and purification of C17/25

At the time this project was started only little information about the C17/25 subcomplex was available, based on biochemical and genetic experiments (Chapter II, Introduction). Based on these data the Pol III subcomplex C17/25 has been suggested to be the counterpart of the subcomplex Rpb4/7 in Pol II (Sadhale & Woychik, 1994; Hu *et al.*, 2002; Siaut *et al.*, 2003).

Previous experimants in our laboratory showed that individually expressed recombinant subunits of the bigger protein complexes are generally insoluble. In fact neither Rpb4 nor Rpb7 could be overexpressed alone in *E.coli* (Armache *et al.*, 2005). Observed insolubility most probably resulted from a loss of structural integrity of the single subunits, which were lacking their natural interaction partner during folding events. For Rpb4/7 this problem was overcome by construction of a double-expression system (Sakurai *et al.*, 1999), and the same strategy was applied for the C17/25 subcomplex. The genes encoding full-length C17 and C25 were amplified from the *S.cerevisiae* genomic DNA and cloned into the pET-21b vector (Figure 2.1). Both, the C17 and C25 coding frames were preceded by separate T7 promoters, the C25 coding frame was followed by a His-Tag and T7 terminator. In this way, both nascent peptide chains emerge in physical proximity upon translation from the ribosomes enabling proper folding.
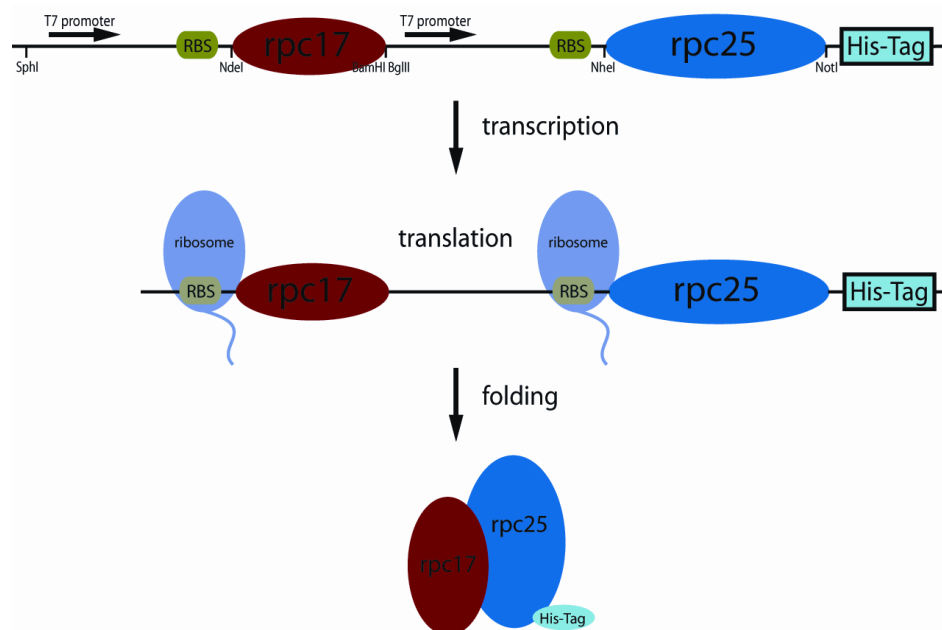
Figure 2.1 Double expression system of C17/25.

The complex was expressed and purified as described in Chapter II, Experimental procedures. First, the *E.coli* lysate was loaded onto a Ni NTA column equilibrated with Lysis buffer containing 1M NaCl. The unspecific binding of the *E.coli* proteins to the chromatographic resin and possible nucleic acids binding to the purified complex could be decreased by the applied high salt concentration, enabeled by a surprising stability of C17/25. The purified complex showed high purity already after the affinity chromatography step and subsequent purification by anion exchange chromatography on a MonoQ column removed remaining contaminants and an excess of the C25 subunit. Gel filtration of the C17/25-containing fractions resulted in a single peak, indicating a stoichiometric protein complex. As a pure and homogeneous material is a prerequisite for crystallization the purity of the peak fractions was confirmed by SDS-PAGE analysis (Figure 2.2). Protein identities were confirmed by mass spectrometry.
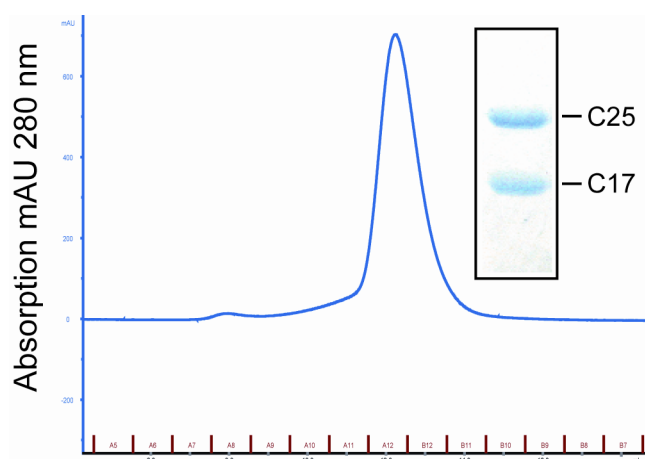
Figure 2.2 Purification of the C17/25 subcomplex.

Chromatogram of the Superose 12 gel filtration. The absorbtion at 280 nm is measured to detect protein elution (blue). SDS-PAGE analysis of the peak fractions is shown next to the peak.

2.2. Limited proteolysis and the protein stability tests

Highly mobile protein regions are known to inhibit crystallization. In order to identify those regions limited proteolysis of the purified C17/25 heterodimer was performed. Proteolysis of the natively folded protein complexes occurs mainly at the highly flexible parts, like loops or solvent exposed partially unfolded domains, which are not involved in protein-protein interactions. Globular or tightly bound domains are rather rigid and therefore more resistant to proteolysis (Fontana *et al.*, 1986; Fontana *et al.*, 2004). Stable fragments obtained as a result of the proteolysis experiment may indicate compact folded regions of the protein complex and hence serve as good candidates for crystallization. The C17/25 subcomplex was incubated with chymotrypsin and trypsin at 37 °C and the reaction was stopped after 10 sek, 1, 3, 10, 30 and 60

minutes and samples were analyzed by SDS-PAGE. Two stable cleavage products were observed (Figure 2.3 a, red boxes) and were analyzed by   EDMAN-sequencing. Both products, starting with the sequences SKGKQ and NVVNY, proofed to be the C-terminal part of the C17. The trypsin cleavage sites were at the positions 41 and 60, respectively, in the amino acid sequence of C17 (Figure 2.7 a, marked with a black arrow).

Stability tests of the C17/25 complex were carried out to test the behaviour of the proteins in the conditions similar to that of the crystallization experiment (solution and 20 ºC). The freshly purified protein solution in gel filtration buffer was incubated at room temperature for more than 4 days. The solution stayed clear, without any traces of protein precipitation and the samples taken after 0, 24, 48 and 96 h show no pro-teolitytic degradation (Figure 2.3 b).
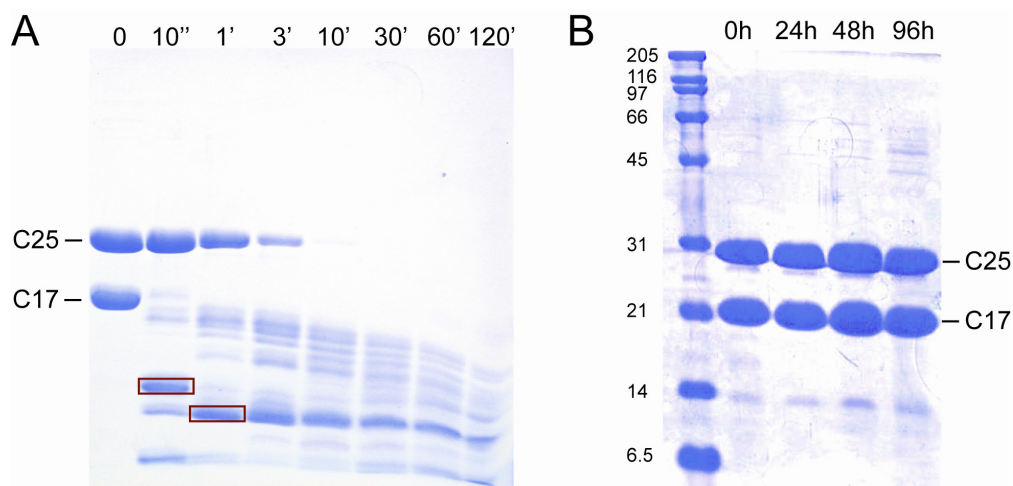
Figure 2.3 SDS-PAGE analysis of the trypsin cleavage (A) and protein stability tests (B) of C17/25.
Incubation time in seconds ('') and minutes (') is indicated above the gels. The names of the proteins as well as molecular weight (kDa) of the marker are given. Red boxes indicate the limited proteolysis products analysed by EDMAN-sequencing.

2.3. Crystalization of C17/25

Parallel to the partial proteolysis experiments and protein stability tests purified samples of C17/25 were subjected to crystallization trials. Initial crystal setups with commercial screens were performed with the Hydra II semi-automatic protein crystallization robot (Matrix Technologies Apogent Discoveries) by sitting drop vapour diffusion methods using 96-well crystallization plates (Corning). Nextal Classic, MPD, Anions and Cations Suite (Nextal/Qiagen), as well as Hampton Index, Natrix and PEG/Ion Screens (Hampton Research) were tested. First screens, with a protein concentration of 4 mg/ml, yielded no crystals. Only repeated screening with a higher protein concetration (8.5 mg/ml) gave rise to microcrystals of the full-length C17/25 in two conditions of the Nextal Classic screen (Figure 2.4 a, b). Both conditions

contained 0.1M Hepes pH 7.5 and either 1.6M ammonium sulphate and 2% PEG 400 or 2M ammonium sulphate and 0.1M NaCl. Further screens were carried out manually using a hanging drop method. Fine screening of the pH, ammonium sulphate concentration and PEG or NaCl concentration, respectively, resulted in obtaining small but well shaped crystals of the C17/25 subcomplex in both solutions. Crystals with the length of 170 µm and diameter 80 µm obtained in 2% PEG 400, 1.4M ammonium sulphate and 0.1M Hepes pH 7.5 were tested at the Swiss Light Source and diffracted up to 4.5 Å (Figure 2.4 c and d). In the final optimization screen the improved conditions were combined resulting in a crystallization solution composed of 0.1M NaCl, 3% PEG 400, 1.6M ammonium sulphate and 0.1M Hepes pH 7.5. The presence of NaCl in the crystallization solution increased protein solubility and reduced a number of nucleization events, which resulted in fewer but larger crystals. Well-shaped protein crystals appeared usually 24 hours after setting up the crystallization experiments and reached their full size (350 µm x 140 µm diameter) during the next three days (Figure 2.4 e, f). Most of them were firmly attached to the cover slide of the crystallization plate well or to the skin covering the crystallization drop. Best crystals were very carefully detached from the plastic or the skin with the help of an acupuncture needle and harvested with the loop. Cryo-cooling was achieved by a step-wise transfer of the crystal to the final cryo solution and by plunging into liquid nitrogen (for the details see Chapter II, 4.15).
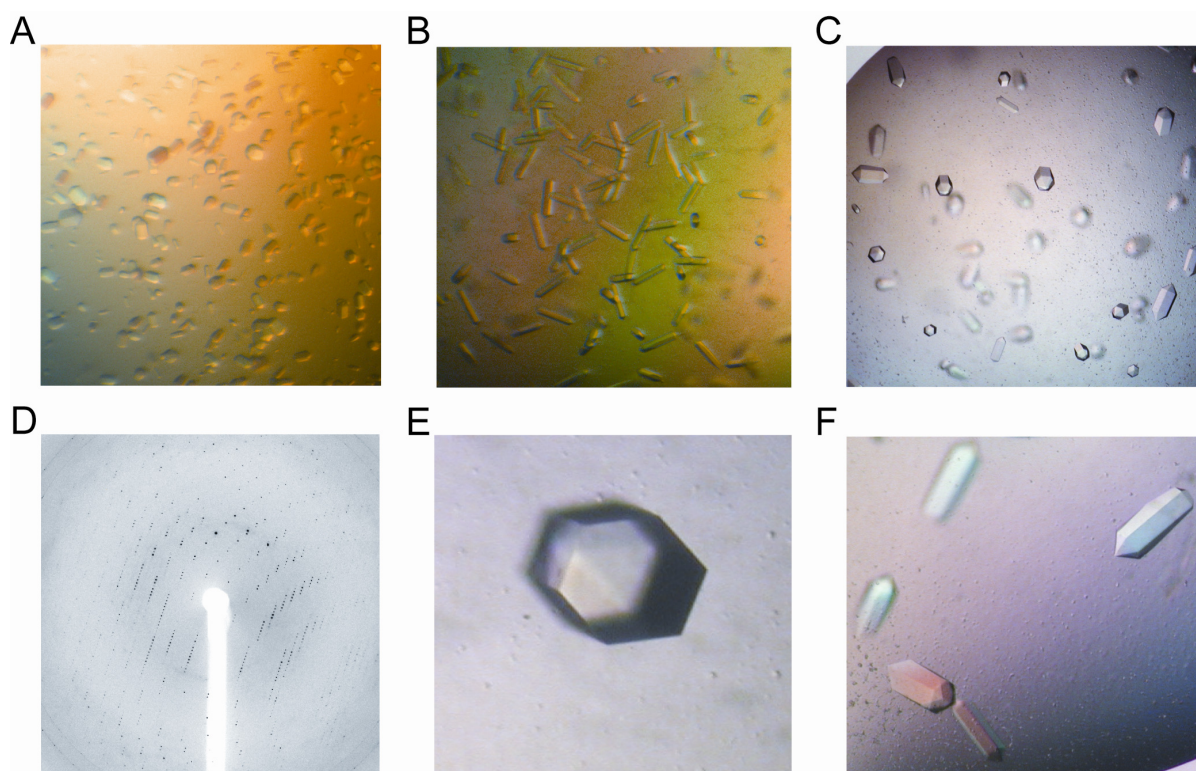
Figure 2.4 Crystallization of C17/25.
A. **B**. Initial microcrystals obtained in the Nextal Classic Screen. **C**. Refined Crystals tested on the SLS.
D. Diffraction pattern of the C17/25 crystals. **E**. **F**. Final optimized crystals of the C17/25 complex.

2.4. Purification and crystallization of SeMet-labeled C17/25

The selenomethionine-labeled C17/25 complex was expressed in the methionine auxotroph *E.coli* strain B834 (DE3). Bacteria were grown in minimal medium supplemented with selenomethionine and antibiotics (Budisa *et al.*, 1995), (Meinhart *et al.*, 2003) and an overnight expression of the protein was induced by the addition of 0.5 mM IPTG to the culture (for the details see Chapter II, 4.7). SeMet-labelled proteins were purified like the native C17/25 subcomplex. SeMet-labelled C17/25 was very well expressed and the high purity of the complex after the size exclusion chromatography step was confirmed by SDS-PAGE analysis (Figure 2.5 a). The peak fractions were concentrated to 8.9 mg/ml, flash-frozen in liquid Nitrogen and stored in -80ºC.
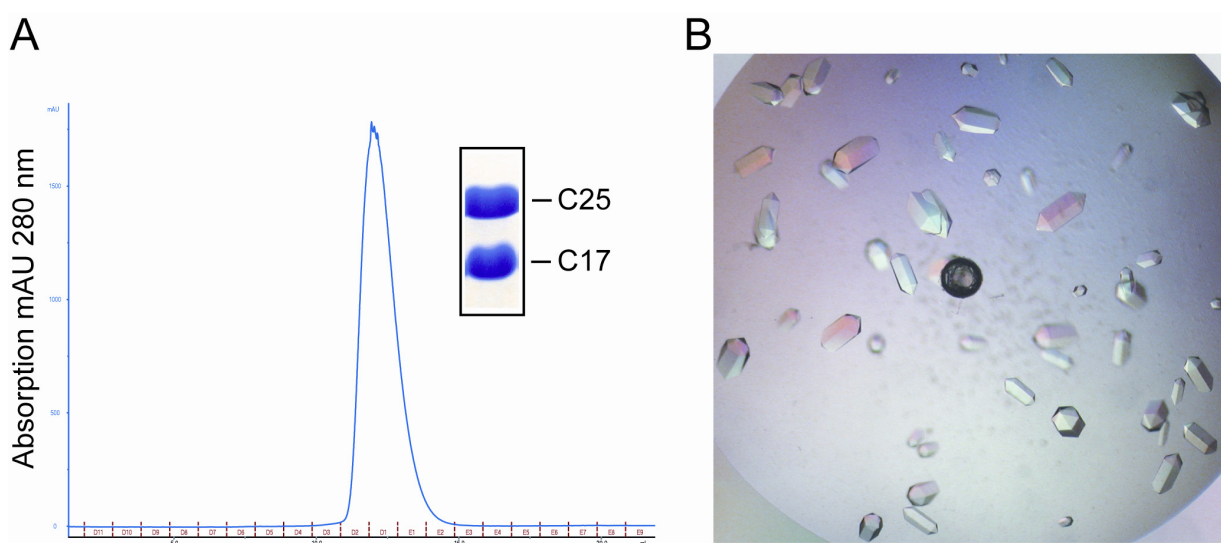
Figure 2.5 SeMet-labeled C17/25.

A. Chromatogram of the Mono Q. The absorbtion at 280 nm is measured to detect protein elution (blue). SDS-PAGE analysis of the peak fractions is shown next to the peak. **B**. Crystals of SeMet C17/25

Crystallization trials of SeMet-labeled C17/25 were performed manually in 24-well plates with the hanging-drop method. Apart from the crystallization in the native C17/25-optimized condition two additional fine screens was carried out. In fine screens, the concentration of Hepes pH 7.5 (0.1M) and PEG 400 (3%) was kept constant, while the concentration of NaCl (50 - 200mM) and ammonium sulphate (1.2 - 1.7M) were varied. In the second screen, PEG 400 (0 - 3%) and ammonium sulphate (1.2 - 1.7M) concentration were varied by a stable concentration of NaCl (0.1M) and Hepes pH 7.5 (0.1M). Best crystals, with 200 μm length and 100 μm diameter, could be obtained in 150mM NaCl, 3% PEG 400, 1.6M ammonium sulphate and 0.1M Hepes pH 7.5 (Figure 2.5 b). Crystals were harvested and cryo-cooled like the native C17/25 crystals (see Chapter II, 4.15).

2.5. X-ray analysis of the Pol III subcomplex C17/25

All diffraction data were collected with an increment of 1 degree per image at the beamline X06SA at the Swiss Light Source, Villigen, Switzerland and processed with DENZO and SCALEPACK (Otwinowski & Minor, 1996). Since molecular replacement with the structures of Rpb4/7 (Armache *et al.*, 2005) and the archaeal counterpart RpoF/E (Todone *et al.*, 2001) failed, the crystals were phased de novo with selenomethionine labeling and single-wavelength anomalous dispersion (Table 2.1). A total of 18 selenium peaks that stemmed from two C17/25 complexes in the asymmetric unit were detected with programs SnB and SOLVE (Weeks & Miller, 1999; Terwilliger, 2002). After SAD phasing with all sites in SOLVE, a model was automatically built by program RESOLVE and manually corrected with the program O (Jones *et al.*, 1991). The native structure was solved by molecular replacement with the obtained model. The structure was refined at 3.2 Å resolution to a free R-factor of 30.7% and shows good stereochemistry (Table 2.1). In the refined structure, 99% of the residues fall in allowed and additionally allowed regions of the Ramachandran plot, and none of the residues are in disallowed regions.

Tabele 2.1 C17/25 X-ray diffraction data and refinement statistics.

Crystal	SeMet C17/25	Native C17/25
Data collection[1]		
Space group	$P6_122$	$P6_122$
Wavelength (Å)	0.97932	0.97894
Unit cell axis (Å)	137.5, 240.6	138.2, 247.1
Resolution (Å)	50–3.5 (3.63–3.5)[2]	30–3.2 (3.31–3.2)
Completeness (%)	100	88.6 (91.4)
Unique reflections	17,684 (1,726)	21,061 (2,107)
Redundancy	14.4 (14.7)	3.9 (3.9)
R_{sym} (%)	10.3 (38.5)	9.1 (46.9)
$<I/\sigma I>$	10.6 (8.3)	17.6 (2.6)
Refinement		
Amino acid residues		537
RMSD bonds (Å)		0.007
RMSD angles (°)		1.3
R_{cryst} (%)		23.6
R_{free} (%)		30.7

[1]Diffraction data were collected at beamline X06SA at the Swiss Light Source, Villigen, Switzerland.
[2]Numbers in parenthesis refer to the highest resolution shell.

2.6. Overall C17/25 structure

The structure of C25 resembles that of its counterparts Rpb7 and RpoE (Figure 2.7). The N-terminal "tip" domain of C25 shows an RMS deviation in Cα atom positions of 4.2 Å and 1.6 Å in Rpb7 and RpoE, respectively, whereas the C-terminal OB domain is quite divergent. The relative position of the two C25 domains differs slightly from that observed in Rpb7 (Figure 2.6 a). C25 differs from Rpb7 mainly by the absence of the short helical turn K* in the tip domain and the presence of a flexible, non-conserved loop B4-B5 that is 34 residue longer than in Rpb7 (Figure 2.7).

The structure of C17 reveals a compact N-terminal "tip-associated" domain, which packs mainly against the C25 tip domain, and not between the tip and OB domains as in Rpb4/7 and RpoF/E (Figure 2.7 b). The only contact between the C17 tip-associated domain and the C25 OB domain is formed between C17 helix H2 and C25 loop B2-B3. Consistently, a mutation at the B2-B3 loop (S100P) impairs C17 binding in vivo (Zaros & Thuriaux, 2005). The C17 tip-associated domain connects via a flexible linker to a C-terminal HRDC domain, a fold that occurs in RecQ helicases and ribonucleases (Morozov *et al.*, 1997; Meka *et al.*, 2003). The C17 HRDC fold resembles the corresponding domains in Rpb4 and RpoF (RMSD in Cα positions of 2.07 and 4.5 Å, respectively, Figure 2.6 b), although the sequence conservation is very weak or absent (Figure 2.7 a, Table 2.2).
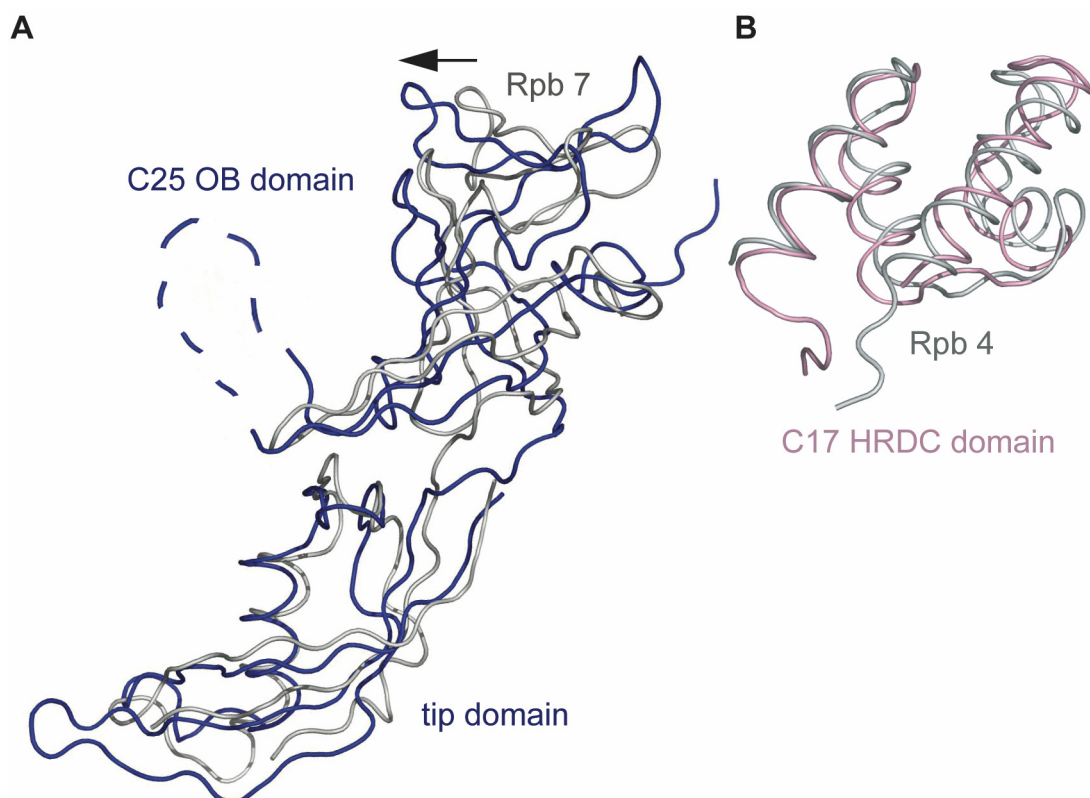
Figure 2.6 Comparison of domain folds in C17/25 and Rpb4/7.

A. Comparison of C25 (blue) and Rpb7 (silver) based on superposition of their tip domains. **B**. Superposition of HRDC domains in C17 (pink) and Rpb4 (silver).
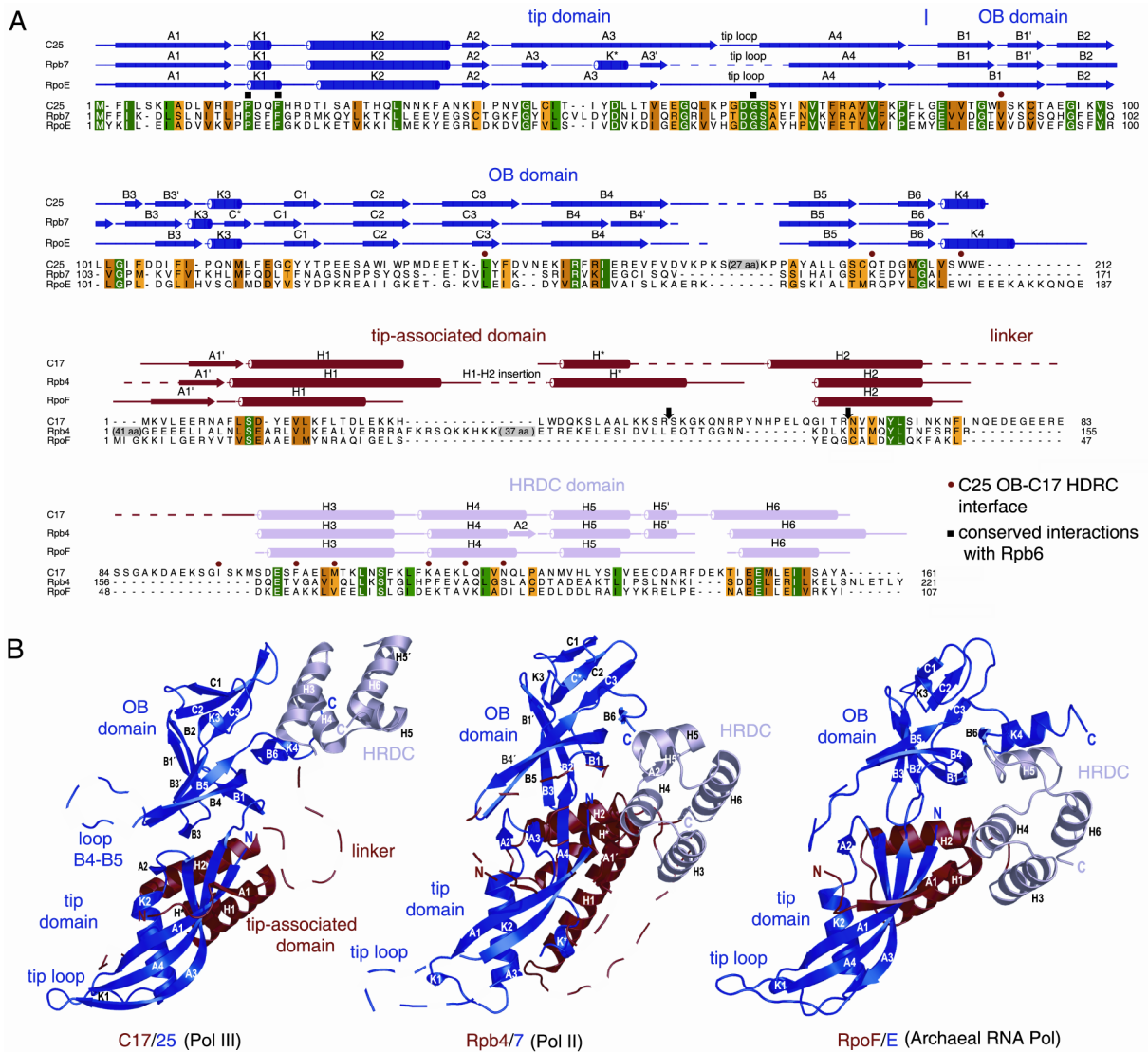
Figure 2.7 X-ray structure of the Pol III subcomplex C17/25.

A. Primary and secondary structure. Structure-based alignments of amino acid sequences of *S. cerevisiae* C25 (top) and C17 (bottom) with their counterparts in Pol II (*S. cerevisiae* Rpb7 and Rpb4, respectively) and archaeal RNA polymerase (*M. jannaschii* RpoE and RpoF, respectively). Secondary structure elements are shown above the sequences (cylinders, α-helices; arrows, β-strands; lines, loops; dashed lines, disordered). Conserved residues are highlighted according to decreasing conservation from green, through orange, to yellow. Cleavage sites revealed by limited proteolysis with trypsin are indicated with arrows. Three C25 residues involved in conserved interactions with the Rpb6 are indicated with a black square. Residues that contribute to the C25-C17 HRDC interface are indicated with a red dot. **B**. Comparison of the structures of yeast C17/25 (this study, left) with that of yeast Rpb4/7 (Armache *et al.*, 2005) (center) and archaeal RpoF/E (Todone *et al.*, 2001) (right). C25/Rpb7/RpoE are in blue and C17/Rpb4/RpoF are in red, with the HRDC domain in light red. Disordered in the C17/25 structure are the C25 loop B4-B5 (residues K59-K90), the C17 loop H*-H2 (residues K38-N47), and the C17 linker between the tip-associated domain and the HRDC domain (residues N69-G94).

2.7. The C17 HRDC domain adopts a unique position

Although the overall domain folds are conserved between C17/25 and Rpb4/7, the observed position of the C17 HRDC domain is very different from that in Rpb4 and RpoF (Figure 2.7 b). Compared to Rpb4 or RpoF, the HRDC domain of C17 is translated by about 35 Å and rotated by about 150°. The C17 HRDC domain packs against the C25 OB domain, between the C2-C3 loop and the C-terminus (Figure 2.7 b, 2.8 a). The HRDC-OB interface is complementary in shape and electrostatics (Figure 2.8 a), and includes many hydrophobic residues, which are well conserved among several species (Figure 2.9), but are generally not conserved in the Pol I and Pol II counterparts (C25 residues F116, W130, M132, L138, and W211, and C17 residues M107, L121, and V124) (Figure 2.8 b). Residues in the Rpb4 HRDC domain-Rpb7 interface are also conserved among *S. cerevisiae, S. pombe*, and human, but not in C17/25. Thus, the C17 HRDC-C25 interface is unique and Pol III-specific, suggesting that the observed position of the HRDC domain is a specific feature of Pol III. The same position of the C17 HRDC domain may in principle be adopted in other species. Although in the human and S. pombe sequences the C17 linker is apparently only three residues long, the distance between the two C17 domains could just be spanned if the residues corresponding to C17 residues 95-100 adopt an extended conformation.
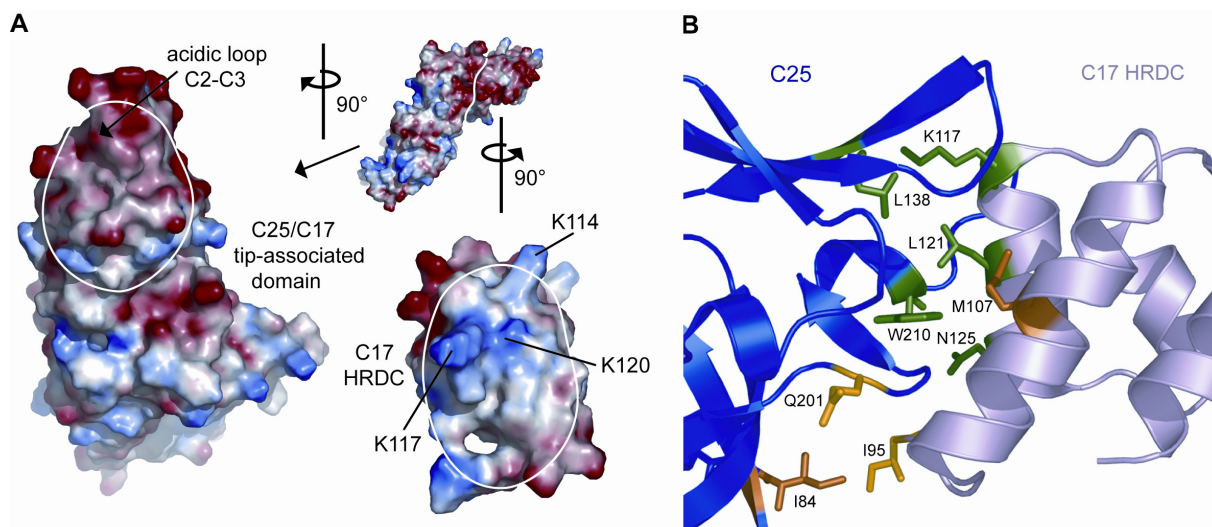
Figure 2.8 Interface between C25 OB domain and the C17 HRDC domain.
A. Electrostatic surfaces. In the center, a surface representation of C17/25 is shown colored according to the electrostatic surface potential (positive, blue; negative, red). The view is as in Figure 2.7 b. A book view of the interface was obtained by separating the C17 HRDC domain from the complex. The remainder of C17/25 and the C17 HRDC domain are depicted to the left and to the right, respectively. **B**. Detailed view of the interface. Residues are highlighted according to decreasing conservation among species from green, through orange, to yellow.
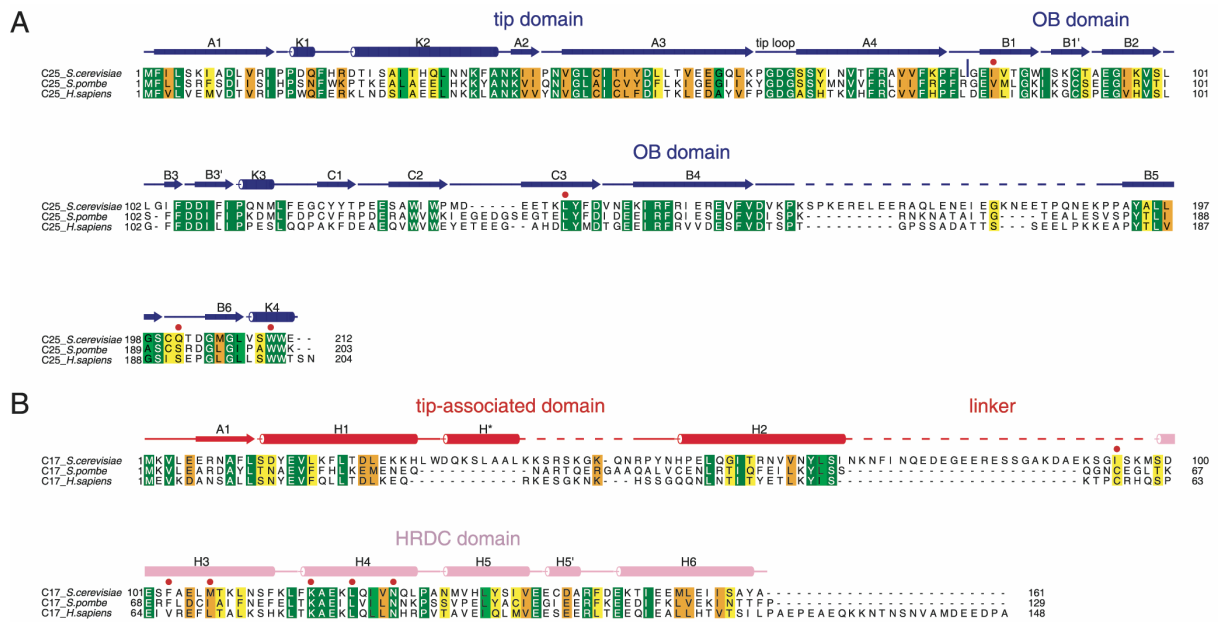
Figure 2.9 Alignment of C25 and C17 from different species.
Conserved residues are highlighted according to decreasing conservation from green, through orange, to yellow. Residues that contribute to the C25-C17 HRDC interface are indicated with a red dot.

2.8. Modular two-domain structure of C17

Although the well-packed nature of the HRDC-OB interface suggests that the location of the HRDC domain is fixed, there is evidence that the domain can change its position. First, the asymmetric unit of the crystals contains two C17/25 heterodimers, but only one HRDC domain is ordered (Figure 2.10 a), whereas the second one is not visible in the electron density, consistent with a low affinity of the HRDC-OB interface. Second, the linker between the C17 tip-associated domain and the HRDC domain is flexible and not conserved among species (Figure 2.9). To test if the C17 HRDC domain forms an independently folding module, the C-terminal part of C17 (amino acids 94 - 161) was cloned into an expression vector (Chapter II, 4.2). The isolated HRDC domain was very well expressed in *E.coli* and could be purified using a combination of affinity, ion exchange and size exclusion chromatography (for details see Chapter II, 4.6 and 4.8.2) (Figure 2.10 b).

Static light scattering revealed that the purified C17 HRDC domain forms a soluble momoneric module which does not tend to aggregate in solution (8.59/8.75 kDa observed/theoretical molecular weight). The extreme stability of the whole complex for several days at room temperature (Figure 2.3) proofs that the flexibility of the HRDC domain does not jeopardize the structural integrity of the C17/25 fold. Together these results show that C17/25 is a modular subcomplex and that the C17 HRDC domain is an autonomously folding module, which may adopt different positions.

Figure 2.10 Modular structure of C17 – HRDC domain.
A. The asymmetric unit of the crystals with two C17/25 heterodimers, but only one ordered HRDC domain. The flexible linker between the C17 tip-associated domain and the HRDC domain is disordered in the crystal structure. The linkers connecting the HRDC domain in the two possible positions are modelled in green. The correct position of the HRDC domain was investigated by HRDC-OB interface analysis (Figure 2.8). C25 is indicated in blue, N-termini of C17 in red and HRDC domain in pink. **B**. SDS-PAGE of the purified HRDC domain. The molecular weight (kDa) of the marker is given.

2.9. Both C17 domains are essential *in vivo*

In contrast to Rpb4, C17 is essential for viability of *S. cerevisiae*. To investigate if the essential *in vivo* function requires both structural domains of C17, I generated plasmids under the control of the heterologous *GAL1* promoter that carried the gene for full-length C17 (*RPC17*), a RPC17 mutant lacking the HRDC domain (*RPC17ΔHRDC*), or the HRDC domain alone. Plasmids were introduced into a *rpc17Δ* yeast strain that was rescued by *RPC17* on a centromeric *URA3* plasmid. Loss of the *URA3* plasmid due to complementation by any of the *GAL1* promoter-driven expression constructs would allow growth on media containing 5'-fluorotic acid (5'-FOA). In these experiments, complementation was allowed by full-length *RPC17*, but not by *RPC17ΔHRDC* or by the HRDC domain alone (Figure 2.11). The same result was obtained at 30° C and 23° C. Thus both C17 domains are required for cell viability and the HRDC domain has an essential function in vivo that requires its proximity to the C17/25 subcomplex.

In an attempt to test if the essential *in vivo* function of the HRDC domain requires its positioning as observed in the crystal structure, I mutated the *RPC17* plasmid such that the hydrophobic residues F103 and M107, both located in the interface between the C17 HRDC domain and the C25 OB domain, were changed to glutamates. The resulting C17 double mutant F103E/M107E is predicted to disrupt the domain interface, but did still support cell growth at either 30° C or 23° C. This supports a possible mobility of the HRDC domain and suggests that the position of the domain in a

functional *in vivo* complex could differ from that observed in the C17/25 crystal structure. It remains however possible that the mutated HRDC domain is retained in the observed position in the context of the complete Pol III enzyme by interactions with additional Pol III subunits.
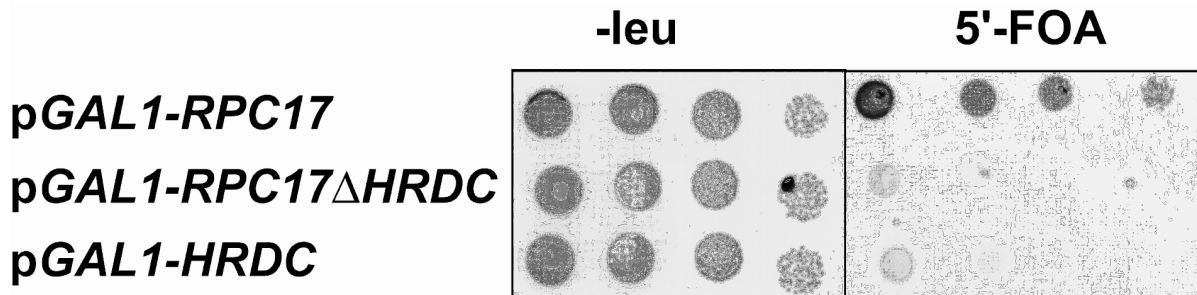
Figure 2.11 The C17 HRDC domain is essential for cell viability.

A yeast strain carrying a *rpc17* deletion and *RPC17* on a centromeric *URA3* plasmid was transformed with centromeric *LEU2* plasmids containing *RPC17*, an *RPC17*ΔHRDC, or the isolated C17 HRDC domain under control of the heterologous *GAL1* promoter. **Left panel**: serial dilutions of each strain grown on -leu plates. **Right panel**: serial dilutions of each strain grown on plates containing 5'-FOA. Only full-length *GAL1-RPC17* can complement for the loss of the *URA3* plasmid carrying wild-type *RPC17* and thus allows growth on 5'-FOA.

2.10. C17/25 binds nucleic acids in vitro

To explore possible functions of C17/25, the purified recombinant subcomplex was subjected to nucleic acid-binding assays (Chapter II, 4.17). I performed non-radioactive electrophoretic mobility shift assays (EMSAs), and revealed the nucleic acids by staining with SYBR-Gold. Similar to Rpb4/7 and Rpa14/43 (Orlicky *et al.*, 2001; Meka *et al.*, 2003; Meka *et al.*, 2005), C17/25 bound to single-stranded RNA (Figure 2.12 a). Comparative EMSA analysis showed that C17/25 bound much stronger to a tRNA sample, with an apparent affinity in the low μM range (Figure 2.12 b). Binding to duplex DNA was also observed, but was weaker than for tRNA (Figure 2.12 c). Single-stranded DNA binding to C17/25 was very weak. All these nucleic acid probes were also bound by recombinant purified Rpb4/7, but generally less efficiently than by C17/25 (Figure 2.12). In particular, Rpb4/7 bound more weakly to tRNA, although it also shows a preference for tRNA relative to the 22 nt ssRNA (Figure 2.12 b). There was generally no indication of sequence specificity in nucleic acid binding. On the longer nucleic acid probes, multiple shifted bands are observed that could correspond to multiple complexes, complicating the interpretation of the results. To check whether the C17 HRDC domain is responsible for the nucleic acid-binding property of C17/25, the EMSA with the isolated HRDC domain were performed. The purified HRDC domain did not show detectable nucleic acid binding, even when a 100-fold or greater excess of the protein over nucleic acids was used (Figure 2.12).

Indeed the HRDC domain surface contains only one small cluster of positively charged residues (lysines K114, K117, K120) that is involved in interaction with the C25 acidic loop C2-C3 (residues D133, E134, E135) (Figure 2.8). In conclusion, there is no evidence for a nucleic acid-binding function of the C17 HRDC domain, but these observations cannot exclude that nucleic acid interactions account for the essential *in vivo* function of the domain.
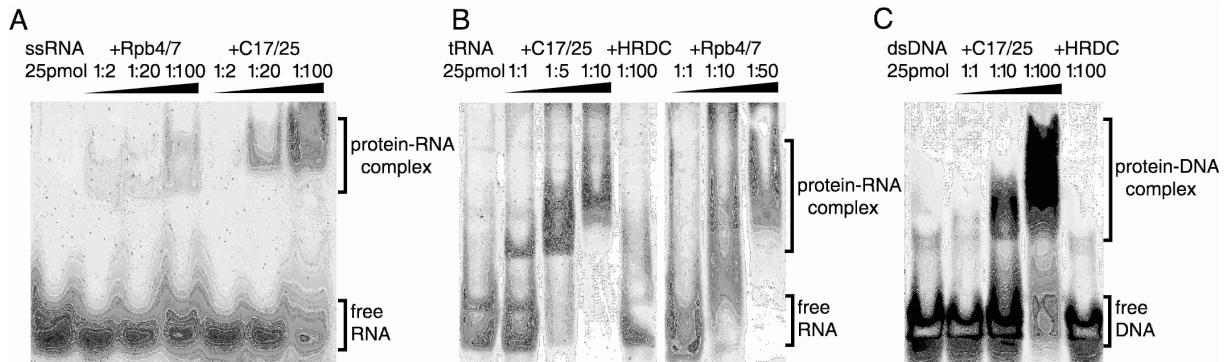
Figure 2.12 Nucleic acid-binding properties of C17/25, the C17 HRDC domain, and Rpb4/7.

A. Comparison of the single-stranded RNA-binding activity of the subcomplexes Rpb4/7 and C17/25. Increasing amounts of the proteins were incubated with 25 pmol of a 22-mer single-stranded RNA and the resulting complexes separated from free RNA by electrophoresis. **B**. C17/25 binds tRNA. Increasing amounts of C17/25 were incubated with 25 pmol of a commercial tRNA preparation from E.coli and the resulting complexes were separated from free tRNA by electrophoresis. The double-bands visible in free and bound tRNA samples may be due to conformational heterogeneity of the tRNA sample. **C**. C17/25 binds duplex DNA. Increasing amounts of C17/25 were incubated with 25 pmol of a 40-base pair duplex DNA and the resulting complexes were separated from free DNA by electrophoresis.

2.11. RNA Pol III model

Generation of the nine-subunit RNA Pol III model was done by Karim Armache and therefore is not directly a part of this thesis. However, the insights into the Pol III architecture gained from combining the Pol III model with the structural information on the C17/25 subcomplex will be presented and discussed here.
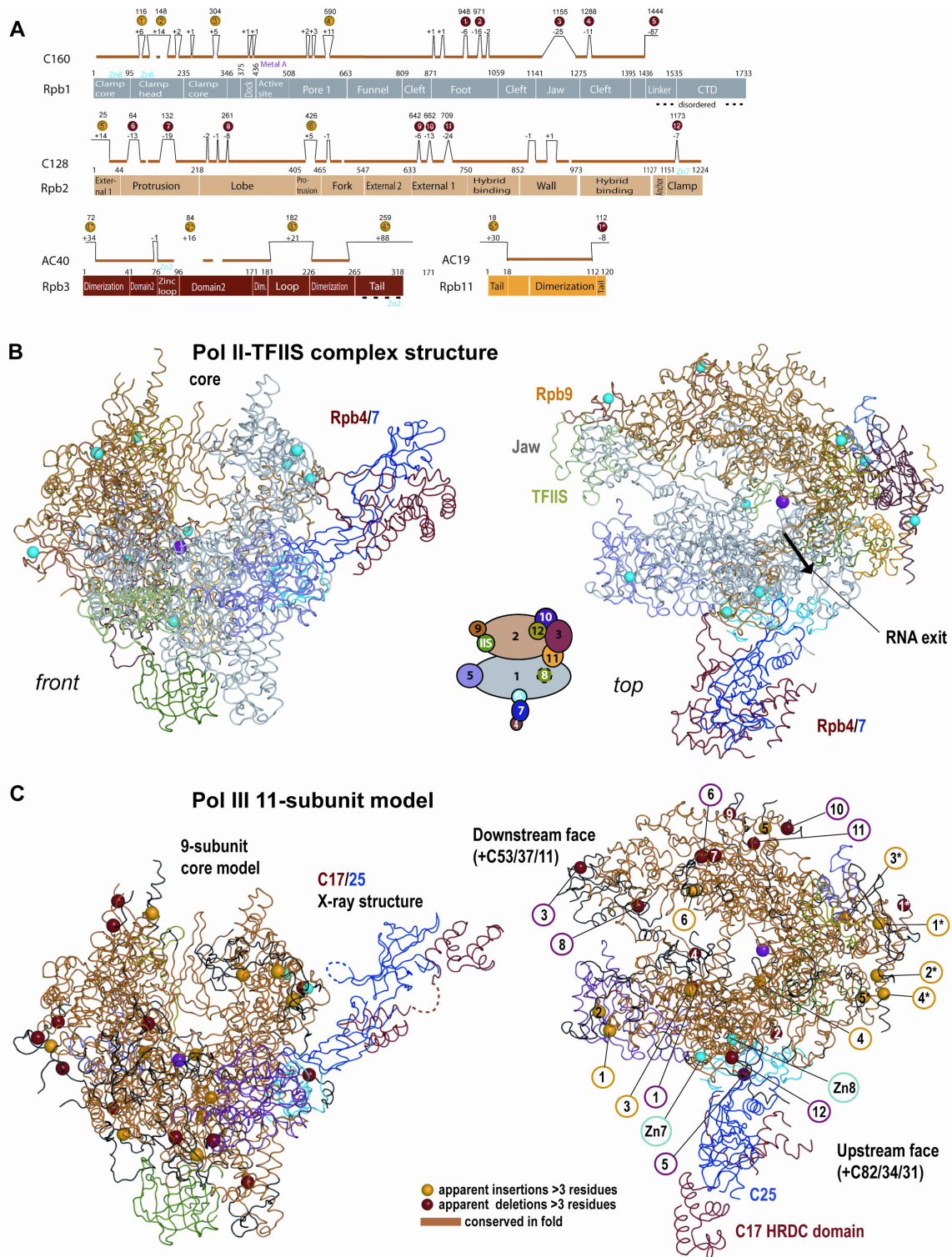
Figure 2.13 11-subunit Pol III model.

A. Pol II structure-guided sequence alignments of core subunits homologous in Pol III. The domain organization of Pol II subunits Rpb1, Rpb2, Rpb3, and Rpb11 is shown as diagrams (Cramer *et al.*, 2001). Above the diagrams, regions conserved in fold in the homologous subunits of Pol III are indicated with orange bars. Regions that apparently adopt a different structure are indicated with black brackets. Indicated are the numbers of residues apparently inserted or deleted in the Pol III subunits (depicted only for more than three residues inserted or deleted). The numbers correspond to the Pol II residue N-terminal (at the beginning) of the insertion/deletion in Pol III. **B**. Structure of the complete Pol II-TFIIS complex (Kettenberger *et al.*, 2003; Kettenberger *et al.*, 2004).

The Pol II subunits Rpb1-Rpb12 and TFIIS are colored according to the key. Eight zinc ions and the active site magnesium ion are depicted as cyan spheres and a pink sphere, respectively. **C**. Model of an 11-subunit form of Pol III. The model was obtained by combining the nine-subunit homology model of the Pol III core with the X-ray structure of the C17/25 complex. Regions in homologous subunits that are conserved in fold are in orange according to **A**. Black regions are not present or altered in Pol III. Deletions and insertions in Pol III amino acid sequences as compared with Pol II are shown as red and yellow spheres, respectively, if they exceed three amino acid residues in length **A**. The two Pol III subcomplexes C82/34/31 and C53/37/11, which lack from the model, are predicted to locate to the upstream and downstream face, respectively.

2.11.1. Interaction of C17/25 with core Pol III

In the Pol II structure, the Rpb4/7 subcomplex binds the Pol II core via two loops in its Rpb7 tip domain, the "tip loop" and the A1-K2 loop (Armache *et al.*, 2003; Bushnell & Kornberg, 2003; Armache *et al.*, 2005). The two corresponding loops in C17/25 are ordered in one copy of the asymmetric unit of our crystals, and adopt a similar conformation as in Pol II. We superimposed the two core-binding loops of C25 with those of Rpb7, to place the C17/25 structure onto the Pol III core model. The resulting C25 tip-core Pol III interface reveals that key contacts with the common core subunit Rpb6 are formed by C25 residues P15, F18, and G64, which are conserved in Rpb7 (Figure 2.7 a), indicating that the placement of C17/25 is correct. The specificity in the interaction between C17/25 and the Pol III core may arise at least partially from a salt bridge between glutamate E56 in C25 and lysine K1432 in C160, which corresponds to isoleucine I1445 in Rpb1.

2.11.2. 11-subunit Pol III model

The correct docking of the C17/25 X-ray structure onto the homology model of the Pol III core resulted in a model for the 11-subunit central part of Pol III, which lacks only the subcomplexes C82/34/31 and C53/37/11 (Figure 2.13). Due to the C17 HRDC domain, which adopts a new position, and due to the slightly different relative orientation of the C25 tip and OB domains, the model shows an orientation of C17/25 relative to the core Pol III that differs from that of Rpb4/7 in Pol II (Figure 2.13). Compared to Rpb4/7, C17/25 protrudes from the Pol III core more towards the upstream face and RNA exit pore (Figure 2.13). The exact orientation of C17/25 however may differ in the complete Pol III, due to the presence of the two additional subcomplexes or due to some remodeling in the interface between the C25 tip and core Pol III.

3.   Discussion

3.1. Structural biology of Pol III

In this thesis first structural information on Pol III was obtained. A homology model of the nine-subunit Pol III core could be combined with the C17/25 structure. Resulting 11-subunit model for Pol III allows interpretation of biochemical and genetic data and the design of mechanistic studies of Pol III transcription.

The X-ray structure of the Pol III subcomplex C17/25 and biochemical data revealed that the C17 HRDC domain forms an independently folded module that adopts an unexpected new position and may be mobile. Complementation analysis in yeast demonstrated an essential function of the C17 HRDC domain *in vivo*. EMSA analysis revealed that C17/25 binds various types of nucleic acids *in vitro*, including tRNA, but that this activity does not reside in the C17 HRDC domain.

3.2. Conserved structure of the Rpb4/7 complexes

Counterparts of Rpb4/7 in yeast RNA Polymerase I (Shpakovski & Shematorova, 1999; Peyroche *et al.*, 2002; Meka *et al.*, 2003) and III (Sadhale & Woychik, 1994; Hu *et al.*, 2002; Siaut *et al.*, 2003) and in the archeal RNA Polymerase (Todone *et al.*, 2001) were identified previously. C17 was identified as the last of the Rpb4 homologues due to a very weak sequence homology between these two subunits (Siaut *et al.*, 2003). Structural information on Rpb4/7 itself and the archeal RpoE/F complex was available (Todone *et al.*, 2001; Armache *et al.*, 2005) and the X-ray structure of the C17/25 heterodimer is part of this thesis (Jasiak *et al.*, 2006). Recently the crystals of the Rpb4/7 counterparts in Pol I could be obtained. Accessibility of the atomic-resolution model of the Rpa14/43 complex allows a comparative analysis of the Rpb4/7-like complexes in all three eukaryotic RNA polymerases (Kuhn *et al.*, 2007) (Figure 3.1).

The overall structure of Rpa14/43 resembles its counterparts Rpb4/7 (Armache *et al.*, 2005), C17/25 (Jasiak *et al.*, 2006), and the archaeal RpoF/E (Todone *et al.*, 2001). Compared to Rpb7, the N-terminal tip domain of Rpa43 shows RMS deviations in $C\alpha$ atom positions of 2.2-2.5 Å, whereas the C-terminal OB domain is more divergent. Rpa14 contains the 'tip-associated domain', formed by two helices H1 and H2 divided by a flexible loop, packs on the Rpa43 tip domain (Kuhn *et al.*, 2007). The most striking feature of the Rpa14 subunit is the lack of an HRDC domain present in all counterparts, instead having a long flexible C-terminal tail (Figure 3.1).

The overall similar fold of Rpb4/7 counterparts in RNA Pol I, II and III provides new insight into the evolution of the eukaryotic RNA polymerases. The conserved structure speaks for a conserved role of Rpb4/7-like complexes in transcription. They are

all crucial for promoter-directed transcription initiation and are shown to interact with Rrn3, TFIIB and TFIIIB, transcription factors recruiting the polymerase to its promoter in the Pol I, II and III system, respectively. The diverse position of the HRDC domain in Rpb4 and C17 and its lack in Rpa14 are likely to be functionally relevant and may influence the differential initiation factor interactions and promoter-specificity of the three polymerases.
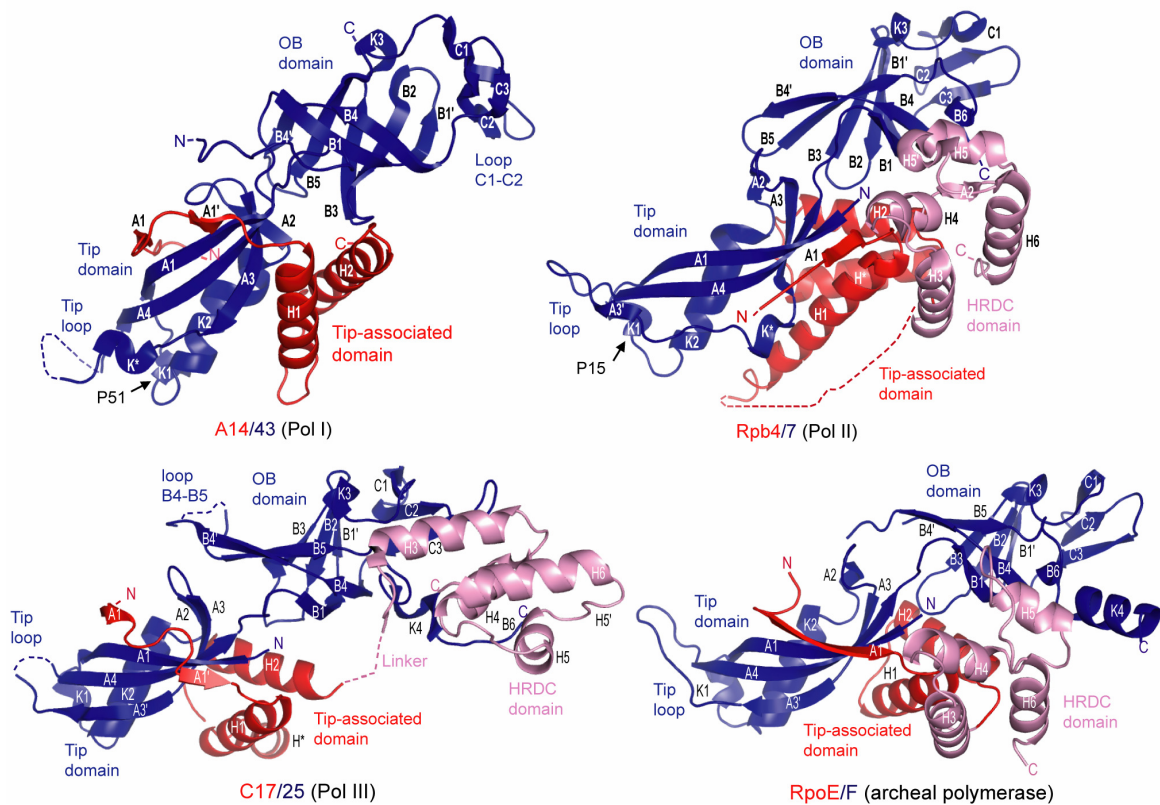
Figure 3.1 Rpb4/7 subcomplexes structures: A14/43, Rpb4/7, C17/25 and RpoE/F. Figure adapted from (Kuhn et al., 2007).

3.3. RNA binding of Rpb4/7-like subcomplexes

The functional conservation of RNA polymerases may extend to binding of the exiting RNA to the Rpb4/7-like subcomplexes, which are located in proximity of the RNA exit pore. Indeed RNA emerging from a Pol II elongation complex can be crosslinked to Rpb7 (Ujvari & Luse, 2006). Further, Rpb4/7, A14/43 and RpoE/F all bind single-stranded nucleic acids (Orlicky et al., 2001; Meka et al., 2003; Meka et al., 2005). C17/25 binding of the single-stranded RNA was showed, suggesting that interaction with exiting RNA is a common property of all Rpb4/7-like complexes. C17/25 bound most strongly to tRNA, suggesting that it may preferentially interact with Pol III tran-scripts that emerge from the RNA exit pore located in the vicinity of the subcomplex (shown for Pol II as a black arrow in Figure 2.13). Consistently, Rpb4/7 bound more weakly to tRNA, and the surfaces of the two subcomplexes are not conserved, in-

cluding an RNA-binding patch of Rpb4/7 (Meka *et al.*, 2005). It is thus possible that emerging Pol III transcripts fold cotranscriptionally on the C17/25 surface or that C17/25 plays a role in coupling Pol III transcription and RNA processing. However, more experiments must be conducted to rigorously address the possibility of preferential interaction of C17/25 with various Pol III transcripts.

3.4.   Promoter-specific initiation

Pol III selects its promoters with specific transcription initiation factors that assemble on the enzyme upstream face. The Pol III upstream face is structurally different from that of Pol II, and includes C17/25 with its differently positioned HRDC domain. The functional importance of C17/25 for initiation is established by a point mutation in C25 that results in an initiation defect (Zaros & Thuriaux, 2005). Indeed C17/25 contributes to initiation complex assembly, since C17 binds to the Pol III initiation factor TFIIIB (Ferri *et al.*, 2000), and to the Pol III subcomplex C82/34/31 (Figure 2.13) (Geiduschek & Kassavetis, 2001; Schramm & Hernandez, 2002) and references therein). The C82/34/31 subcomplex also binds the Pol III core near a zinc site (Werner *et al.*, 1992) (Zn8 in Figure 2.13 c), where the Pol III subunit C128 shows a specific seven-residue deletion (Figure 2.13, deletion 12). C82/34/31 bridges to the initiation factors TFIIIC (Hsieh *et al.*, 1999) and TFIIIB (Geiduschek & Kassavetis, 2001; Schramm & Hernandez, 2002) and references therein). Since C17 is part of an extensive protein interaction network, and since the C17 HRDC domain has an essential *in vivo* function but apparently lacks nucleic acid-binding activity, the HRDC domain is most likely involved in Pol III initiation complex assembly. C17/25 could however also contribute directly to promoter binding since it has an affinity for duplex DNA. In this respect it is interesting that Rpb4/7 can be crosslinked to promoter DNA in a Pol II initiation complex (Chen *et al.*, 2004).

3.5. Polymerase conservation and elongation

Our 11-subunit RNA Pol III model reveals that at least 83.4 % of the Pol II fold is conserved in Pol III, although the overall sequence identity is only 39.4 % (Table 3.1). The only Pol II domain folds that are not present or strongly altered in Pol III are the Rpb1 jaw domain and the Rpb2 external domain 1 (Cramer *et al.*, 2001) (Figure 2.13). Other differences concern only insertions and deletions on the enzyme surface (Figure 2.13). Comparison with the Pol II elongation complex structures (Gnatt *et al.*, 2001; Kettenberger *et al.*, 2004; Westover *et al.*, 2004) shows that extended surfaces conserved between Pol III and Pol II are only found in the polymerase cleft, around the incoming DNA, in the active site, around the binding sites for the nucleoside triphosphate substrate and the DNA-RNA hybrid, and at the RNA exit tunnel (Figure

3.2), reflecting the conservation of the basic mechanisms of RNA elongation in the two polymerases.
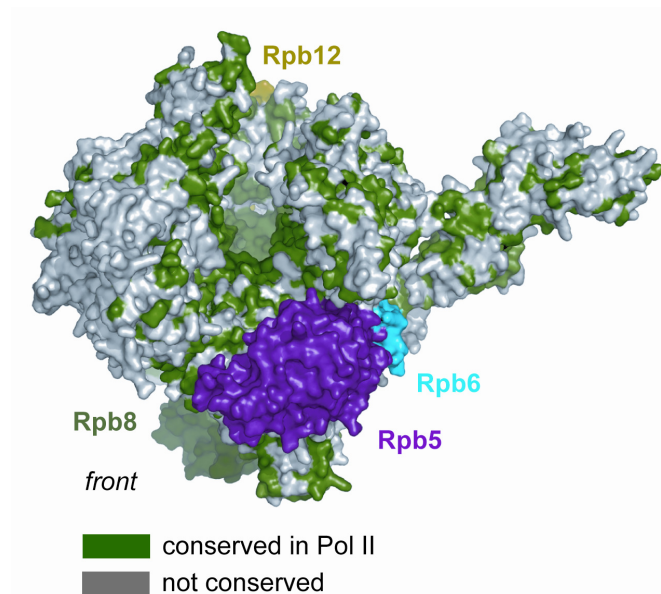
Figure 3.2 Surfaces on the Pol III model conserved in Pol II.

Residues on the surface of the 11 subunit Pol III model that are identical or conserved in the four homologous subunits of *S.cerevisiae* Pol II are in green, others are in silver. The common subunits are in different colors. Extensive conservation is observed only in the active center cleft.

Table 3.1 Conservation of the RNA polymerase III subunits.

Polymerase part	Pol III subunit	Pol II subunit	Subunit type	Sequence identity (%)	Conserved fold (%)
Core	C160	Rpb1	homolog	28.4	83.2
	C128	Rpb2	homolog	35.8	87.2
	AC40	Rpb3	homolog	25.8	60.2
	AC19	Rpb11	homolog	20.8	81.6
	Rpb5	Rpb5	common	100	100
	Rpb6	Rpb6	common	100	100
	Rpb8	Rpb8	common	100	100
	Rpb10	Rpb10	common	100	100
	Rpb12	Rpb12	common	100	100
Rpb4/7 subcomplex	C17	Rpb4	homolog	7.2	50
	C25	Rpb7	homolog	25.2	81.3
Upstream subcomplex	C82/34/31		specific	-	-
Downstream subcomplex	C53/37/11*	Rpb9*	unclear	-	-
11-subunit Pol III model	-	-	-	39.4	83.4

*Subunit C11 shows homology to Rpb9 and TFIIS. Rpb9 was previously defined as a part of the Pol II core.

3.6. Evaluation of the RNA Pol III model

Two recent publications inspecting RNA Polymerase III architecture using mass spectrometry (Lorenzen *et al.*, 2007) and cryo-electron microscopy (Fernandez-Tornero *et al.*, 2007) could evaluate the structural Pol III model presented here. They are both consistent with our results, which proofs suitability of the structure-based modeling combined with the X-ray analysis method to analyze big multi-subunit protein complexes.

DMSO treatment of purified Pol III coupled to tandem mass spectrometry measurements allowed for dissociation and identification of the peripheral Pol III subcomplexes (Lorenzen *et al.,* 2007). According to the obtained data, Pol III compromises a 10 subunit core that resembles the structure of the Pol II core, and three heterodimeric subcomplexes C37/53, C17/25 and C34/82. C31 bridges between C82/34, C17/25 and the Pol III core. These observations support our results and expand the model of Pol III by identifying the C11 subunit, which could not be modeled, as a part of the Pol III core. The EM structure of Pol III at 17 Å resolution revealed a hand-like shape typical of RNA polymerases (Fernandez-Tornero *et al.*, 2007). Difference map calculated after fitting of the RNA Pol II structure into the obtained electron density revealed several additional and a few smaller missing density patches (Figure 3.3). The three largest additional densities were assigned to the C37/53 (AD2) and C31/34/82 (AD1 and AD3) complexes, other differential densities co-localize with the deletions and insertions in Pol III compared to Pol II, which were marked on our model (Figure 2.13).
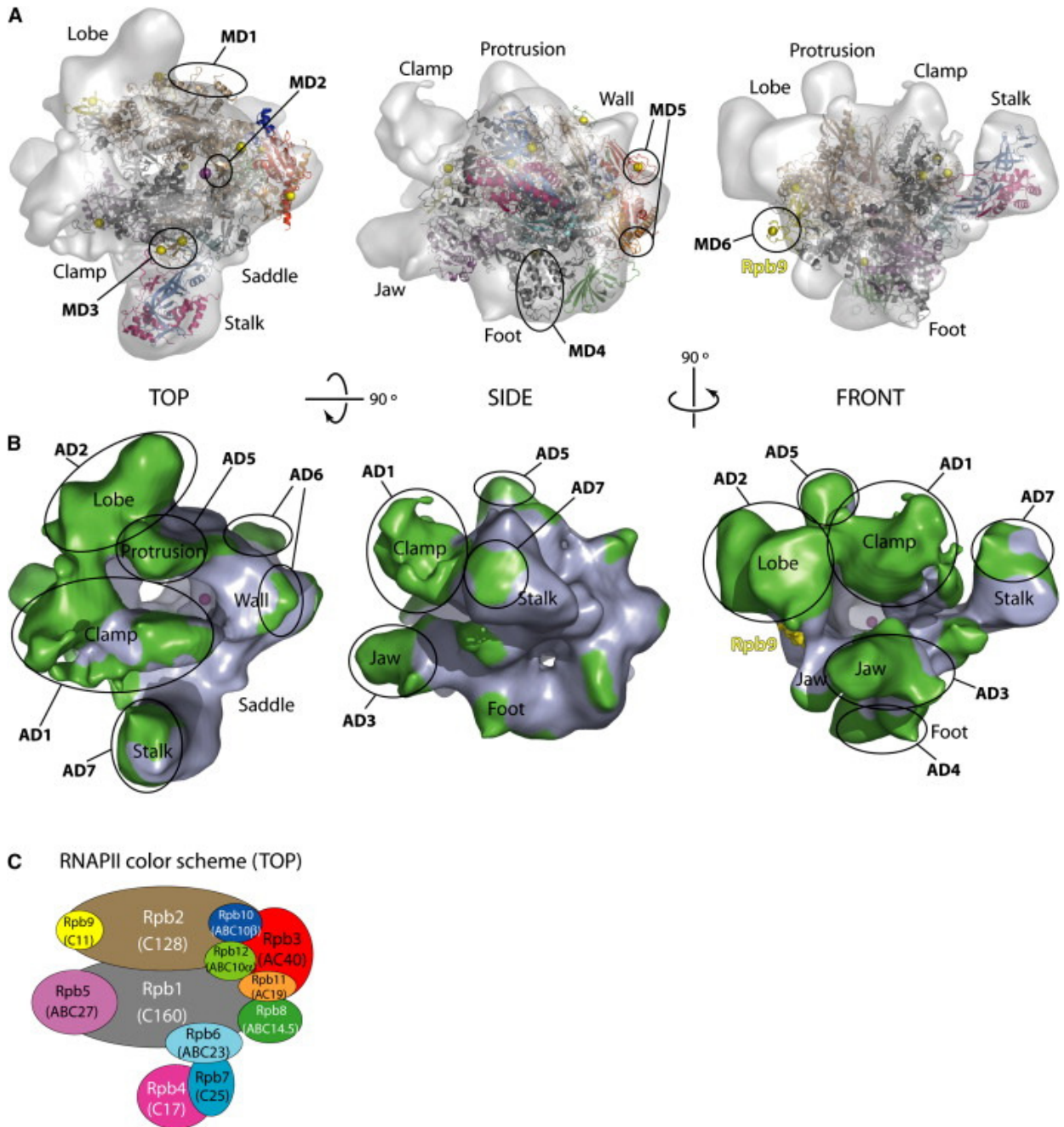
Figure 3.3 Structurel differences between RNA Pol III and RNA Pol II. Figure adapted from (Fernandez-Tornero *et al.*, 2007).

A. Final manual fitting of RNAPII crystal structure (Armache *et al.*, 2005) into the reconstructed RNAPIII density. Domains of RNAPII that fall outside the EM density are labeled MD1–MD6. Mg^{2+} and Zn^{2+} ions appear as magenta- and yellow-colored dots, respectively. **B**. Difference map between the RNAPIII cryo-EM structure and the RNAPII crystal structure. Additional density features (green) in the RNAPIII reconstruction are labeled AD1–AD7. The N-terminal domain of Rpb9, positioned according to the fitting in (A), is in yellow. The putative position of the active center is marked with a magenta dot. **C**. Schematic representation and color code of RNAPII subunits, with corresponding RNAPIII subunits in parenthesis.

3.7. Mobility of the HRDC domain

The EM structure of the Pol III confirms that C17/25 adopts the same fold in the context of the whole Pol III enzyme as it does, when crystallized alone (Fernandez-Tornero *et al.*, 2007). The observed additional density surrounding both subunits (Figure 3.4) could be attributed to a linker connecting the N- and C-terminal domains of C17 and a loop in the OB domain of C25, which are disordered in the crystal structure of the complex (this work). Moreover superposition of the C17/25 crystal structure onto Rpb4/7 in the Pol II structure fitted into the Pol III electron density resulted in the HRDC domain protruding almost entirely from out of the EM map (Figure 3.4 a). In contrast, positioning the C17 HRDC domain as observed in Rpb4 perfectly fits the density (Figure 3.4 b). This and the notion that the largest difference between the native and negatively stained EM structures is observed in the C17/25 region and might reflect a different position of the C-terminus of C17 (Fernandez-Tornero *et al.*, 2007) support demonstrated in this thesis mobility of the HRDC domain.
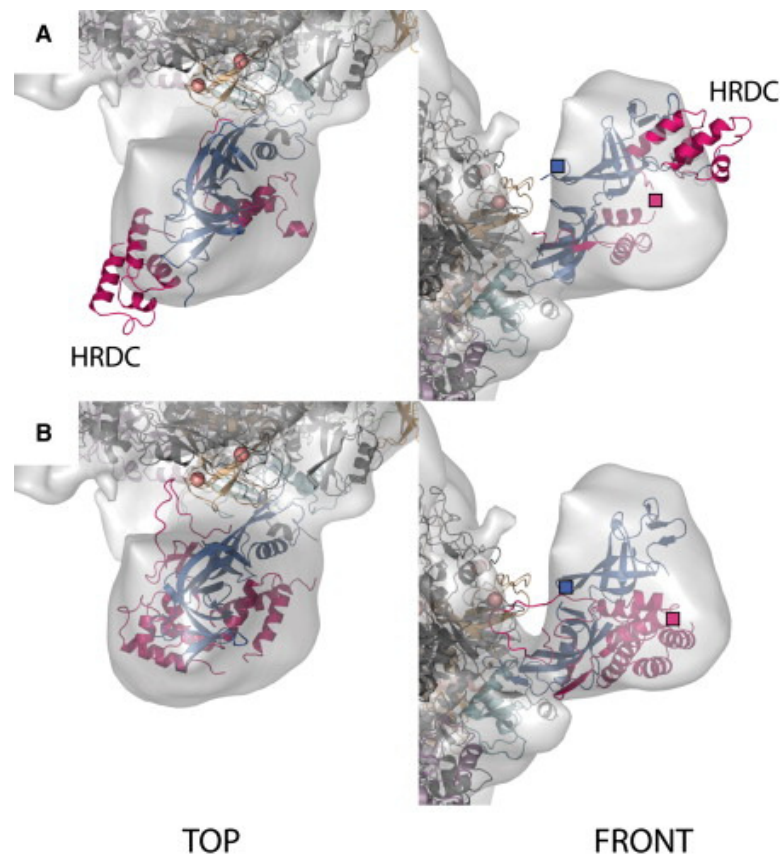
Figure 3.4 Close-up view of the C17/25 region in the RNA Pol III EM structure. Figure adapted from (Fernandez-Tornero *et al.*, 2007).

The RNA Pol II model fitted into the EM reconstruction **A**. Subunits Rpb4/Rpb7 were replaced by the RNAPIII C17/C25 subcomplex by superposing subunit C25 onto Rpb7. Squares indicate the position of two insertions in C17/C25 that are disordered in the crystal structure. **B**. Subunits Rpb4/Rpb7 in the Pol II model. Squares indicate the position of two insertions in the RNAPIII subunits C17/C25.

3.8. Conclusions and outlook

Pol III is a key enzyme for the expression of the eukaryotic genome, which for a very long time resisted structural investigation. Here, a combination of X-ray crystallography, molecular modeling, and functional *in vitro* and *in vivo* analysis, was used to establish an 11-subunit model of Pol III, and to provide molecular insights into similarities and differences with Pol II.

Whereas the similarity of the core fold and active center of Pol III and Pol II reflects a common basic transcription mechanism, structural differences in a region that directs initiation complex assembly partially account for promoter specificity. These data provide a framework for further structural and functional analysis of Pol III, and mark the beginning of a comparative molecular analysis of eukaryotic RNA polymerases. Further work on crystallization and functional analysis of the other Pol III subcomplexes is required to explain the additional densities in the EM map or to solve the X-ray structure of the whole Pol III enzyme in the future. A search for the structural and functional similarities with Pol II transcription factors could enlighten the specific role played by the additional subunits in the relatively simple Pol III transcription machinery.

4. Experimental procedures

4.1. Molecular biology methods

Standart molecular biology procedures like Standard molecular biology procedures such as isolation of DNA, restriction analysis, polymerase chain reaction (PCR), cloning of DNA, agarose gel electrophoresis etc. were carried out essentially as described bevore (Sambrook & Russell, 2001).

4.2. Cloning

The genes for C17 and C25 were amplified from yeast genomic DNA by PCR and were cloned into vector pET21b (Novagen), using double-expression system previously described for rpb 4/7 (Sakurai *et al.*, 1999), resulting in a C-terminal hexahistidine tag (His) on C25. For cloning of the HRDC domain, used for the nucleic acids binding experiments, the region of the gene corresponding to amino acid residues 94-161 of C17 was amplified by PCR from the C17/25 expression plasmid and cloned into pET21b (Novagene) with a C-terminal His tag. Table 4.1 summarises oligonucleotides used for PCR amplification.

Table 4.2 lists all the vectors used in the experiments.

Table 4.1 Oligonucleotides used for PCR

name	5'→3' sequence	restriction site
rpc25_for	ATGCTAGCTAGCATGTTTATCCTTTCAAAAATCGC	NheI
rpc25_rev	ATAGTTTAGCGGCCGCTTCCCACCAACTAACGAGTCCC	NotI
rpc17_for	GGGAATTCCATATGAAAGTTCTTGAGGAAAGG	NdeI
rpc17_rev	CGCGGATCCTCATGCGTACGCAGAGATGATC	BamHI
HRDC_for	GGGAATTCCATATGGGTATAAGCAAAATGAGCG	NdeI
HRDC_rev	CCGGGCGAGCTCTCAGTGGTGGTGGTGGTGGTGTGCGTA CGCAGAGATGATC	XhoI

Table 4.2 Plasmids

name	protein	vector	restriction site	affinity Tag
pET_C17/25_fl	C17	pET21b	NdeI/BamHI	-
	C25		NheI/NotI	His$_6$
pET_HRDC	HRDC domain of C17	pET21b	NdeI/XhoI	His$_6$

4.3. Bacterial strains

Bacterial strains used for cloning and expression of recombinant proteins are listed in Table 4.3 .

Table 4.3 Bacterial strains

strain	description	source
XL-1 blue	recA1 endA1 gyrA96 thi⁻1 hsdR17 supE44 relA1 lac [F'proAB lacI^{q}ZΔM15 Tn10 (Tet^r)]	Stratagene
BL21(DE3) RIL	*E. coli* B F⁻ *ompT hsdS*(r$_B$⁻ m$_B$⁻) *dcm+* Tet^r *gal* λ (DE3) *endA* Hte [*argU ileY leuW* Cam^r]	Stratagene
B834 (DE3)	F⁻ *ompT hsdS*$_B$ (r$_B$⁻ m$_B$⁻) *gal dcm met* (DE3)	(Budisa *et al.*, 1995)

4.4. Media and buffers

The double-distilled water was used in all the recipies. All media and buffers were autoclaved or sterile filtered for storing and usage.

Luria Bertani medium

1 %	tryptone
0.5 %	yeast extract
0.5 %	NaCl
	pH 7.0

1.5–2 % (w/v) of Bacto-Agar was added to the medium to prepare LB-agar plates.

Auto-inducing Medium: ZY medium

10 g	tryptone

5 g	yeast extract
925 ml	H_2O

Auto-inducing Meduim: 20 x NPS buffer

66 g	$(NH_4)_2SO_4$
136 g	KH_2PO_4
142 g	Na_2HPO_4
900 ml	H_2O

Auto-inducing Meduim: 50 x 5052 solution

250 g	glycerol
25 g	glucose
100 g	α-lactose
730 ml	H_2O

Auto-inducing Meduim: $MgSO_4$ solution

1 M	$MgSO_4$

Minimal Medium

7.5 mM	ammonium sulfate
8.5 mM	NaCl
55 mM	KH_2PO_4
100 mM	K_2HPO_4
1mM	$MgSO_4$
20mM	glucose
1 µg/l	Trace elements (Cu^{2+}, Mn^{2+}, Zn^{2+}, Mo_4^{2-})
10 mg/l	Thiamine
10 mg/l	Biotine
1 mg/l	Ca^{2+}
1 mg/l	Fe^{2+}
100 mg/l	amino acids (A, C, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, W, Y)
100 mg/l	selenomethionine

Supplements/Antibiotics

100 mg/ml in H2O	ampicillin
50 mg/ml in ethanol	chloramphenicol
1 M in H2O	IPTG

100 X proteaseinhibitor mix/ethanol

3 mg/l	leupeptin
14 mg/l	pepstatin A
1.7 g/l	PMSF
3.3 g/l	benzamidine

Edman buffer

| 200 mM | Tris·Cl pH 8.5 |
| 2 % | SDS |

4.5. Transformation

A 50 µl aliquot of competent cells and plasmid DNA were thawed on ice. 1 µl of plasmid DNA was added to BL21(DE3) RIL or B834 (DE3) cells, which were incubated on ice for 15 minutes, followed by a 'heat shock' performed at 42 °C for 60 seconds. After short cooling on ice and adding 1ml of LB medium, the cells were incubated at 37 °C for 45 minutes. The cells were finally plated on LB agar plates containing ampicillin (100 mg/ml) and chloramphenicol (50 mg/ml). The plates were incubated overnight at 37 °C.

4.6. Expression of recombinant proteins in E.coli

A fresh colony was picked within 24 h of transformation, and a starter culture was grown over-night in 50 ml LB medium containing ampicillin and chloramphenicol. LB medium (two liter per five liter flask) supplemented with antibiotics was inoculated and shaken (180 rpm) at 37 °C until the culture reached log phase (OD 0.7 - 0.8). After cooling the *E. coli* suspension on ice IPTG at a 1:2000 ratio was added and the cells were grown over night at 18 °C. Cells were collected by centrifugation (4000 rpm, SLS6000 rotor) at 4 °C, subsequently suspended in Lysis Buffer, flash-frozen in liquid nitrogen and stored at -80 °C.

For expression of HRDC so called Auto-inducing Medium (Studier, 2005) was used. A starter culture was grown over-night like in standard LB medium. For preparing 1

liter of Auto-indcing Medium about 925 ml of ZY medium was complemented with 50 ml of 20 x NPS buffer, 20 ml 50 x 5052 solution and 1 ml 1M MgSO$_4$. Auto-indcing Medium (two liter per five liter flask) supplemented with antibiotics was inoculated and shaken (180 rpm) at 30 °C for about 3 – 4 h, until the culture reached log phase (OD 0.7 - 0.8), then the temperature was shifted to 18 °C and the cells were grown over nigh. Cells were collected by centrifugation (4000 rpm, SLS6000 rotor) at 4 °C, subsequently suspended in HDRC Lysis Buffer, flash-frozen in liquid nitrogen and stored at -80 °C.

4.7. Seleno-Methionine labelling

For seleno-methionine labelling pET21b C17/25 vector was transformed into the methionine auxotroph E.coli strain B834 (DE3). Bacteria were grown in LB medium supplemented with the appropriate antibiotic at 37 °C to an OD 0.5. Cells were harvested and resuspended in the same amount of minimal medium supplemented with selenomethionine (100 mg/l) and antibiotics (Budisa *et al.*, 1995), (Meinhart *et al.*, 2003). Cells were grown until the OD increased by 0.2 at 37 °C to deplete the medium of any residual methionine. Cultures were cooled on ice for 30 min and protein expression was induced by the addition of 0.5 mM IPTG. Proteins were expressed over night at 18 °C.

4.8. Protein purification

4.8.1. Purification of C17/25 heterodimer

4.8.1.1. Buffers

Lysis Buffer

1 M	NaCl
50 mM	Tris·Cl pH 8.0
5 %	Glycerol
10 mM	β-Mercapto-ethanol

A0 Buffer

50 mM	Tris·Cl pH 8.0
1 mM	EDTA
5 mM	DTT

A1 Buffer

150 mM	NaCl
50 mM	Tris·Cl pH 8.0
1 mM	EDTA
5 mM	DTT

B1 Buffer

1 M	NaCl
50 mM	Tris·Cl pH 8.0
1 mM	EDTA
5 mM	DTT

Gel filtration Buffer

40 mM	$(NH_4)_2SO_4$
5 mM	HEPES pH 8.5
10 mM	DTT

4.8.1.2. Purification procedures

Cells were harvested by centrifugation, resuspended in Lysis buffer and lysed by sonication. After centrifugation the supernatant was loaded onto a Ni-NTA column (Qiagen) equilibrated with Lysis buffer. The column was washed stepwise with 10 mL of Lysis buffer and 5 mL of Lysis buffer containing 20 mM imidazole. Proteins were eluted with Lysis buffer containing 150 mM imidazole. Eluted fractions were diluted 6.5-fold with Buffer A0 and further purified by anion exchange chromatography (MonoQ, Amersham). The column was equilibrated with Buffer A1 and proteins were eluted with a linear gradient of 10 column volumes from 150 mM to 1 M NaCl, created by addidtion of the B1 Buffer. After concentration the sample was applied to a Superose-12 HR gel filtration column (Amersham) equilibrated with the Gel filtration buffer. Pooled peak fractions were concentrated for crystallization to 8.5 mg/mL.

4.8.2. Purification of HDRC domain

4.8.2.1. Buffers

HRDC Lysis Buffer

150 mM	NaCl
50 mM	Tris·Cl pH 7.5
5 %	Glycerol
10 mM	β-Mercapto-ethanol

A1 Buffer

150 mM	NaCl
50 mM	Tris·Cl pH 7.5
1 mM	EDTA
5 mM	DTT

B1 Buffer

1 M	NaCl
50 mM	Tris·Cl pH 7.5
1 mM	EDTA
5 mM	DTT

Gel filtration Buffer

40 mM	$(NH_4)_2SO_4$
5 mM	HEPES pH 7.5
10 mM	DTT

HDRC domain of C17, used for nucleic acids binding experiments, was purified essentially as described for full length C17/25 complex, except lowering the salt concentration in the HRDC Lysis Buffer. The purified HRDC domain was concentrated to 10 mg/mL, flash-frozen in liquid nitrogen, and stored in -80° C.

4.9. Measurement of protein concentration

For the determination of protein concentrations the Bradford protein assay was used (Bradfor, 1976). Dye reagent was purchased from Biorad and the assay was performed according to the manufacturer's instructions. For each new batch of dye reagent a calibration curve was generated using Bovine serum albumin (Fraktion V, Roth).

4.10. Protein separation by SDS-PAGE

Denaturing gel electrophoresis was adapted to separate complex protein mixtures into distinct bands on a gel. According to the discontinuous Laemmli system (Sambrook & Russell, 2001), ten gels were cast at once. The percentage of the gels (15 % -17 %) was defined by the size of the monitored proteins. The proteins were totally unfolded by adding β-mercaptoethanol to the SDS loading dye. Gelswere then stained with Coomassie (SIGMA) solution and if required subjected to blotting procedures.

4.11. Limited proteolysis

For chymotrypsin and trypsin treatment 1 µg of the protease was added to 20 µg to 50 µg of purified protein. Digests were done in the buffers used for gel filtration and supplemented with $CaCl_2$ to a final concentration of 4 µM. The mixture was incubated at 37 °C and aliquots were removed at 10 sek, 1, 3, 10, 30 and 60 minutes. The reactions were stopped by the addition of SDS sample buffer and were heated immediately to 95 °C for 5 min. All samples were analyzed by SDS-PAGE. Bands of interest were passively transfered to PVDF membrane and analyzed by Edman sequencing.

4.12. Blotting and Edman Sequencing

For N-terminal sequencing proteins were separated on SDS-PAGE and stained with Coomassie. After electrophoresis, the desired band of interest was excised and dried in a Speed Vac. Further the piece of gel was rehydrated in 50 µl of Edman buffer and 100 µl of distilled water was added to set up a concentration gradient together with a small piece of pre wet (ethanol) PVDF membrane. Once the solution began to turn blue from dye, methanol was added to a final concentration of 10 %. Protein transfer required two days incubation at room temperature, after which the solution became clear and the membrane piece blue, indicating completing of the procedure. The

membrane was washed 5 times with 10 % Methanol by vortexing 30 sec each time, dried and the strip was loaded into a PROCISE 491 sequencer (Applied Biosystems).

4.13.   Static light scattering experiments

Static light scattering measurements were performed with a triple detector TDA (viscotek) connected to a Superose-12 HR gel filtration column (Amersham) equilibrated with gel filtration buffer. 100 µl of protein with a concentration of 10 mg/ml was diluted to 500 µl with gel filtration buffer and injected onto the column. The concentration/volume was calculated from the UV absorption and extinction coefcient. With this information the refractive index/volume was calculated by means of the refractive index detector. Refractive index, UV and viscosity were followed during the measurement. The hydrodynamic radius and the molecular weight were calculated using the static light scattering software from viscotek with the calculated refractive index.

4.14.   Temperature stability tests

To test the stability of the purified C17/25 complex 100 µl of newly purified protein solution in gel filtration buffer was incubated in room temperature for over 4 days. The samples were taken after 0, 24, 48 and 96 h, mixed with SDS sample buffer and heated to 95 °C for 5 min. All samples were analyzed by SDS-PAGE.

4.15.   Protein crystallization

Initial crystal setups with commercial screens were performed with the Hydra II semi-automatic protein crystallization robot (Matrix Technologies Apogent Discoveries) by sitting drop vapour diffusion methods using 96-well crystallization plates (Corning). 0.5 µl protein and 0.5 µl crystallization solution drops with 50 µl reservoir solution were set. For manual refinement of the initial crystallization conditions 24 well plates EasyXtal Tools (Nextal/Qiagen) and hanging-drop method were applied. Drops were set using 500 µl reservoir solution and 1 µl protein + 1 µl reservoir drops. In all cases protein was added prior to adding reservoir solution, which always contained fresh reducing agent, 3 mM Tris(2-carboxyethyl)phosphine (TCEP). Crystals of native C17/25 subcomplex were grown at 20°C, using as reservoir solution 3% PEG400, 0.1 M Hepes pH 7.5, 1.6 M ammonium sulphate, and 0.1 M NaCl. For SAD phasing, selenomethionine-containing C17/25 was crystallized as above, except that 0.15 M NaCl was used in the reservoir solution.

Crystals were harvested in reservoir solution containing additionally 5% glycerol. The glycerol concentration was increased to 30% in four steps, using an incubation time of 2 min per step. The crystals were flash-cooled by plunging into liquid nitrogen.

4.16. X-ray structure analysis

The SAD experiment on selenomethionine-labeled crystals and native crystal measurements were performed at the Swiss Light Source at X06SA beamline. Data were processed with DENZO and SCALEPACK (Otwinowski & Minor, 1996). The program SnB (Weeks & Miller, 1999) detected 12 selenium sites, which were used as seeds in program SOLVE (Terwilliger, 2002). This resulted in a total of 18 selenium peaks that stemmed from two C17/25 complexes in the asymmetric unit. After SAD phasing with all sites in SOLVE, program RESOLVE was used for density modification and NCS averaging, and was able to auto-build an initial model, which was corrected and completed manually with program O (Jones *et al.*, 1991). Since native crystals showed slightly different unit cell dimensions, the resulting model was used for molecular replacement with program PHASER (McCoy *et al.*, 2005). The C17/25 model was subsequently refined against native data at 3.2 Å resolution to a free R-factor of 30.7%. In the refined structure, 99% of the residues fall in allowed and additionally allowed regions of the Ramachandran plot, and none of the residues are in disallowed regions. Figures were prepared with Pymol (DeLano, 2002).

4.17. Nucleic acids binding assay

4.17.1. Gels and Buffers

Binding Buffer 1 x

40 mM	HEPES pH 7.2
100 mM	K Acetate
10 %	Glycerol

TBE 0.5 x

45 mM	Tris - Borate
1 mM	EDTA

7.5 % polyacrylamide (PAA) gels (20 ml – 2 gels)

2 ml	5 x TBE
5 ml	30 % PAA
0.4 ml	Glycerol
130 µl	10 % APS
10 µl	TEMED
12.5 ml	H_2O

4.17.2. Electrophoretic mobility shift assay

Nucleic acids binding activities of C17/25 complex and HRDC domain of C17 in comparision with Rbp4/7 were analyzed using the electrophoretic mobility shift assay (EMSA). Complexes of protein and DNA migrate through a native polyacrylamide (PAA) gel more slowly than free oligonucleotides. C17/25 and HRDC domain purification was performed as described in Chapter II, 4.8, Rpb4/7 complex was prepared as described (Armache *et al.*, 2003).

Nucleic acid probes included single-stranded DNA and RNA 22-mers, a double-stranded 40 base-pair DNA, and an E.coli tRNA preparation (Sigma). Double-stranded DNA probes ware obtained by annealing complementary synthetic single strands. 25 pmol of tRNA, duplex DNA, or ssRNA, or 50 pmol of ssDNA, were incubated with different amounts of protein in 20 µl of binding buffer for 30 min on ice as described (Orlicky *et al.*, 2001). Bound and free probes were resolved by electrophoresis in native 4-20% acrylamide TBE gradient gels (Invitrogen) or in 7.5% polyacrylamide 0.5 x TBE gels containing 2% glycerol, in case of single-stranded nucleic acids, at 100 V at 4° C for 2.5 hr or 40 min, recpectively. Gels were stained for 30 min with SYBR-Gold (Molecular Probes) and visualized on a Typhoon 9400 phosphoimager (Amersham). Afterwards the gels were restained with Coommasie Brilliant Blue R (Roth) solution to localize the protein bands. This ensured co-localization of nucleo-protein complexes.

Table 4.4 Oligonucleotides used for binding assay

name	5'→3' sequence
ssDNA	ATGAAAGTTCTTGAGGAAAGG
ssRNA	UAUAUGCAUAAAGACCAGGC
dsDNA_for	ACCGAAAGCTTTATATAGGCTATTGCCCAAAAATGTATCGCCAATCACCTAATTTGGAG
dsDNA_rev	CTCCAAATTAGGTGATTGGCGCATACATTTTTGGGCAATAGCCTATATAAAGCTTTCGGT

4.18. Yeast complementation studies

4.18.1. Cloning

PCR-amplified ORF of complete *RPC17* or of complete *RPC17,* the N-terminal 80 amino acids of *RPC17* and its C-terminal 68-residue HRDC domain were cloned into p416-*URA3*-p*GAL* or p415-*LEU2*-p*GAL1* vectors, respectively. Start and stop codons were introduced by PCR mutagenesis. Table 4.5 lists primers used for PCR amplification.

Table 4.5 Oligonucleotides used for PCR

name	5'→3' sequence	restriction site
y_C17fl_for	GGCCGCGGATCCATGAAAGTTCTTGAGGAAAGG	BamHI
y_C17fl_rev	GCCGGCGAGCTCTCATGCGTACGCAGAGATGATC	XhoI
y_C17delHRDC_rev	GCCGGCGAGCTCTCACTCGCCTTCATCTTCCTGG	XhoI
y_HRDC_for	GGCCGCGGATCCATGGGTATAAGCAAAATGAGCG	BamHI

4.18.2. Mutagenesis

The double mutation F103E/M107E of C17 was introduced into plasmid p415-*LEU2*-p*GAL1-RPC17* with the use of the QuickChange kit (Stratagene) and following the manufacturer's instructions. Kit used a simplified method to perform point mutations, change amino acids or delete/insert amino acids using a thermal cycling technique in combination with Dpn I digestion. Procedure begined with gene of interest in a double strand vector, purified using a standard plasmid-prep kit (Qiagen). Two primers, both containing the desired mutation and covering the area where the mutation is to be made, were designed according to a detailed set of instructions

provided with the kit. After PCR step with a double strand vector used as an templete, the reaction was cooled below 37ºC and 1 ul of Dpn I restriction enzyme was added to each reaction and incubated at 37ºC for 1hr. Dpn I enzyme is specific to methylated and hemi-methylated DNA. As most E.coli strains produce methylated DNA and are not resistant to Dpn I digestion, it digests the parental DNA template but not the mutant-synthesized DNA. Undigested DNA was then used to transform competent cells. Introduction of the point mutations was further proved by DNA sequencing. Sequences of primers used for mutagenesis are listed in Table 4.6 .

Table 4.6 Oligonucleotides used for mutagenesis

name	5'→3' sequence
C17Mut103FE107ME_for	GAGCGATGAAAGCGAAGCTGAGTTGGAGACTAAACTGAATTC
C17Mut103FE107ME_rev	GAATTCAGTTTAGTCTCCAACTCAGCTTCGCTTTCATCGCTC

4.18.3. Vectors

Plasmids used for complementation of the yeast strains are listed in Table 4.7 .

Table 4.7 Plasmids used for yeast complemantation studies

name	gene	vector	restriction site
RJP1207	RPC17	p416-URA3-pGAL1	BamHI/XhoI
RJP1208	RPC17	p415-LEU2-pGAL1	BamHI/XhoI
RJP1209	RPC17ΔHRDC	p415-LEU2-pGAL1	BamHI/XhoI
RJP1210	HRDC	p415-LEU2-pGAL1	BamHI/XhoI
RJP1208a	*RPC17* Mut F103E/M107E	p415-LEU2-pGAL1	BamHI/XhoI

4.18.4. Yeast complementation

Yeast complementation studies were performed by Birgit Märtens from Ralf-Peter Jansen Group, Gene Center Munich as a part of collaboration.

Diploid S. cerevisiae strain Y26779 containing a heterozygous deletion of the C17 gene *RPC17* (Euroscarf, Frankfurt, Germany) was transformed with plasmid RJP1207 and sporulated. Haploid cells were selected carrying the *RPC17Δ::kan*[R] deletion and the complementing plasmid. The resulting strain (RJY2771) was transformed with plasmids RJP1208, RJP1209, RJP1210 or RJP1208a. Transformed yeast strains were selected on -leu plates containing 2% galactose to induce

expression of the corresponding construct. Single clones were resuspendend and spotted in serial dilutions on plates containg 5'-FOA to test for loss of plasmid RJP1207 indicating complementation by the corresponding construct.

4.19. Bioinformatic tools and software

Alignments. Multiple homologues sequence alignments were initially constructed using CLUSTAL-W (www.ebi.ac.uk/clustalw).

Homology searches. Homologoues sequences were found using the NCBI BLAST and PSI BLAST-Server (http://www.ncbi.nlm.nih.gov/BLAST/)

Secondary structure prediction. Secondary structure prediction was done using the PredictProtein (http://cubic/bioc/columbia/edu/predictprotein.hmtl) secondary structure prediction surfer.

Calculation of molecular weight, absorption coefficient and PI. Calculation of properties of the proteins which are important for the design of the purification strategy as the PI were determined using ProtParam (www.expasy.org/tools/protparam.html). The absorption coefficients and molecular weights used for quantification were obtained form the same server.

Chapter III: Genome-wide distribution of RNA polymerase II and its Rpb4/7 subcomplex in *S. cerevisiae*

1. Introduction

1.1. RNA Polymerase II transcription cycle

Generation of the mRNA by RNA Pol II involve multiple steps (reviewed in (Hahn, 2004), (Cramer, 2007)). Transcription starts by positioning RNA Pol II on the promoter in an inactive form termed the preinitiation complex (PIC) (Bushnell *et al.*, 2004; Miller & Hahn, 2006; Chen *et al.*, 2007). Concomitant transition into the open complex and allow the synthesis of a first phosphodiester bond of RNA. It is characterized by conformational changes in the DNA structure, during which 11-15 bp surrounding the transcription start side are melted and the template strand slips inside the cleft and binds in the Pol II active center. Recognition of the promoter sequences, as well as formation of the preinitiation and subsequently of an open complex is facilitated by several different Pol II specific transcription factors and coactivators (Table 1.1). Other factors are responsible for the chromatin modification in order to allow the polymerase to move along the transcribed gene (Armstrong, 2007; Kulaeva *et al.*, 2007; Li *et al.*, 2007). Release of the promoter-assembled transcription factors and entering of the elongation phase occurs after synthesis of about 30 nucleotides (Figure 1.1). Transcription factors remain bound to the DNA as a scaffold complex marking the transcription start site for re-initiation events. The mRNA production by the elongating Pol II is controlled by several protein factors and coupled to the subsequent events of mRNA biogenesis, like cleavage, capping, polyadenylation and nuclear export. The C-terminal domain of Rpb1 (CTD), with its specific phosphorylation pattern, plays a key role in recruiting factors coordinating the transcription cycle (Hirose & Manley, 2000; Buratowski, 2003; Meinhart *et al.*, 2005). Finally, guided by DNA sequence at the end of the gene, Pol II terminates the transcription by dissociating from the template DNA and release of the synthesized mRNA (Gilmour & Fan, 2008). With the concomitant polymerase recycling and binding to the scaffold complex in the reinitiation event, new round of the transcription cycle can occur.

Table 1.1 *Saccharomyces cerevisiae* **Pol II general transcription factors and coactivators**

Factor		number of subunits	Function (Hahn, 2004)
TFIIA		2	Stabilizes TBP and TFIID-DNA binding. Blocks transcription inhibitors. Positive and negative gene regulation.
TFIIB		1	Binds TBP, Pol II and promoter DNA. Helps fix transcription start site.
TFIID	TBP	1	Binds TATA element and deforms promoter DNA. Platform for assembly of TFIIB, TFIIA and TAFs.
	TAFS	14	Binds INR and DPE promoter elements. Target of regulatory factors.
Mediator		24	Binds cooperatively with Pol II. Kinase and acetyltransferase activity. Stimulates basal and activated transcription. Target of regulatory factors.
TFIIF		3	Binds Pol II and is involved in Pol II recruitment to PIC and in open complex formation.
TFIIE		2	Binds promoter near transcription start. May help open or stabilize the transcription bubble in the open complex.
TFIIH		10	Functions in transcription and DNA repair. Kinase and two helicase activities. Essential for open complex formation. Mutations in IIH can cause human disease.
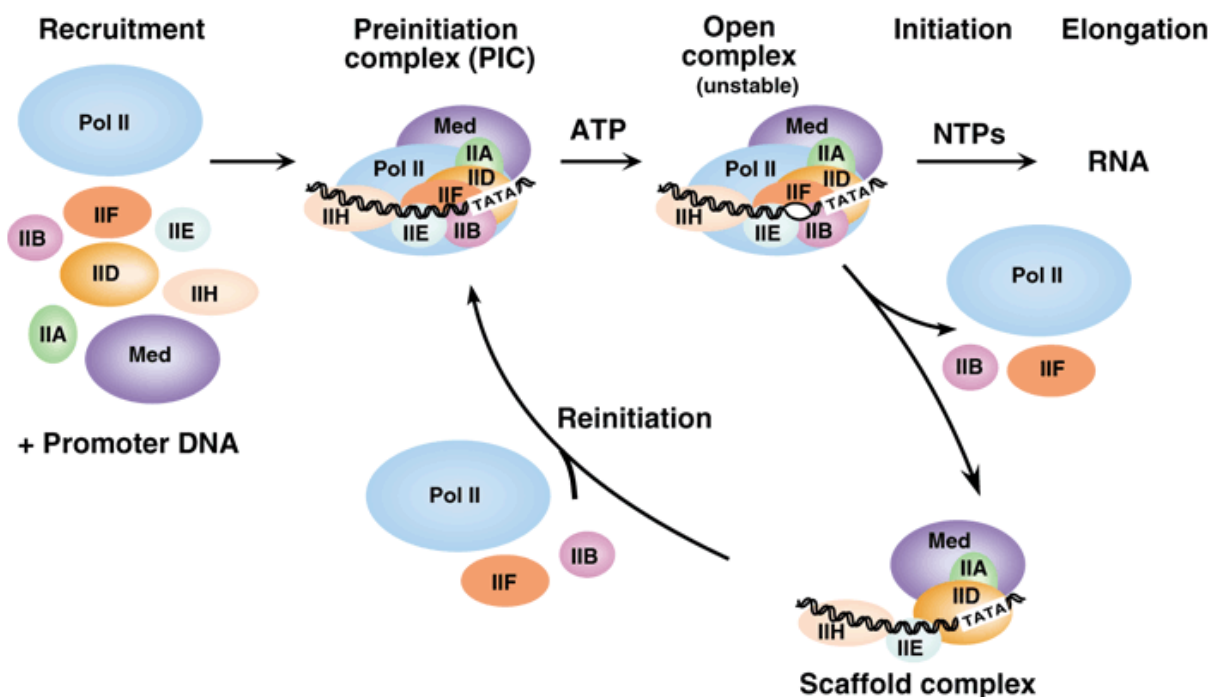
Figure 1.1 The pathway of transcription initiation and reinitiation for Pol II. Adapted from (Hahn, 2004).

See Table 1.1 for a description of each transcription factor and Mediator (Med).

1.2. The RNA Polymerase II structure

Pol II consists of a 10-subunit core enzyme and a peripheral heterodimer of subunits Rpb4 and Rpb7 (Figure 1.2 a). The core enzyme comprises subunits Rpb1, Rpb2, Rpb3, and Rpb11, which contain regions of sequence and structural similarity in Pol I, Pol III, bacterial RNA polymerases (Zhang *et al.*, 1999; Vassylyev *et al.*, 2002), and the archaeal RNA polymerase (Hirata *et al.*, 2008; Kusser *et al.*, 2008). The Pol II core also comprises subunits Rpb5, Rpb6, Rpb8, Rpb10, and Rpb12, which are shared between Pol I, II, and III (common subunits, Chapter I, Table 1). Counterparts of these common subunits except Rpb8 exist in the archaeal polymerase, but only a counterpart of Rpb6 exists in the bacterial enzyme (Minakhin *et al.*, 2001). Finally, homologues of the core subunit Rpb9 exist in Pol I and Pol III, but not in the archaeal or bacterial enzyme. Initial electron microscopic studies of Pol II revealed the overall shape of the enzyme (Darst *et al.*, 1991). The core Pol II could subsequently be crystallized, leading to an electron density map at 6 Å resolution (Fu *et al.*, 1999). Crystal improvement by controlled shrinkage and phasing at 3 Å resolution resulted in a backbone model of the Pol II core (Cramer *et al.*, 2000). This revealed that Rpb1 and Rpb2 form opposite sides of a positively charged active center cleft, whereas the smaller subunits are arrayed around the periphery. Refined atomic structures of the core Pol II at 3.1 and 2.8 Å resolution were obtained in two different conformations and revealed domain-like regions within the subunits, as well as surface elements predicted to have functional roles (Cramer *et al.*, 2001). The active site and the bridge helix, which spans the cleft, line a pore in the floor of the cleft. The Rpb1 side of the cleft forms a mobile clamp, which was trapped in two different open states in the free core structures (Cramer *et al.*, 2001) but was closed in the structure of a core complex that included DNA and RNA (Gnatt *et al.*, 2001). The mobile clamp is connected to the body of the polymerase by five switch regions that show conformational variability. The Rpb2 side of the cleft consists of the lobe and protrusion domains. Rpb2 also forms a protein wall that blocks the end of the cleft. The Pol II core structures lacked subunits Rpb4 and Rpb7, which can dissociate from the yeast enzyme (Edwards *et al.*, 1991). The approximate location of Rpb4/7 on the core polymerase was first determined by electron microscopy (EM) of two-dimensional crystals (Jensen *et al.*, 1998). Later, EM analysis of single particles revealed a closed clamp and showed that the Rpb4/7 subcomplex protrudes from outside the core enzyme below the clamp (Craighead *et al.*, 2002). A different open-closed transition that involved the polymerase jaws was observed by EM of two-dimensional crystals (Asturias *et al.*, 1997). Crystallographic backbone models of the complete 12-subunit Pol II at ≈4 Å resolution revealed the exact position and orientation of Rpb4/7 and showed that it formed a wedge between the clamp and the linker to the unique tail-like C-terminal repeat domain (CTD) of the polymerase

(Armache *et al.*, 2003; Bushnell & Kornberg, 2003). In all crystal structures of the complete Pol II the clamp was observed in a closed conformation. The crystal structure of free Rpb4/7 together with an improved resolution of the complete Pol II crystals finally enabled refinement of a complete atomic model of Pol II (Armache *et al.*, 2005) (Figure 1.2 a). The CTD of Pol II is flexibly linked to the core enzyme and consists of heptapeptide repeats of the consensus sequence YSPTSPS. It is disordered in crystal structures of Pol II.
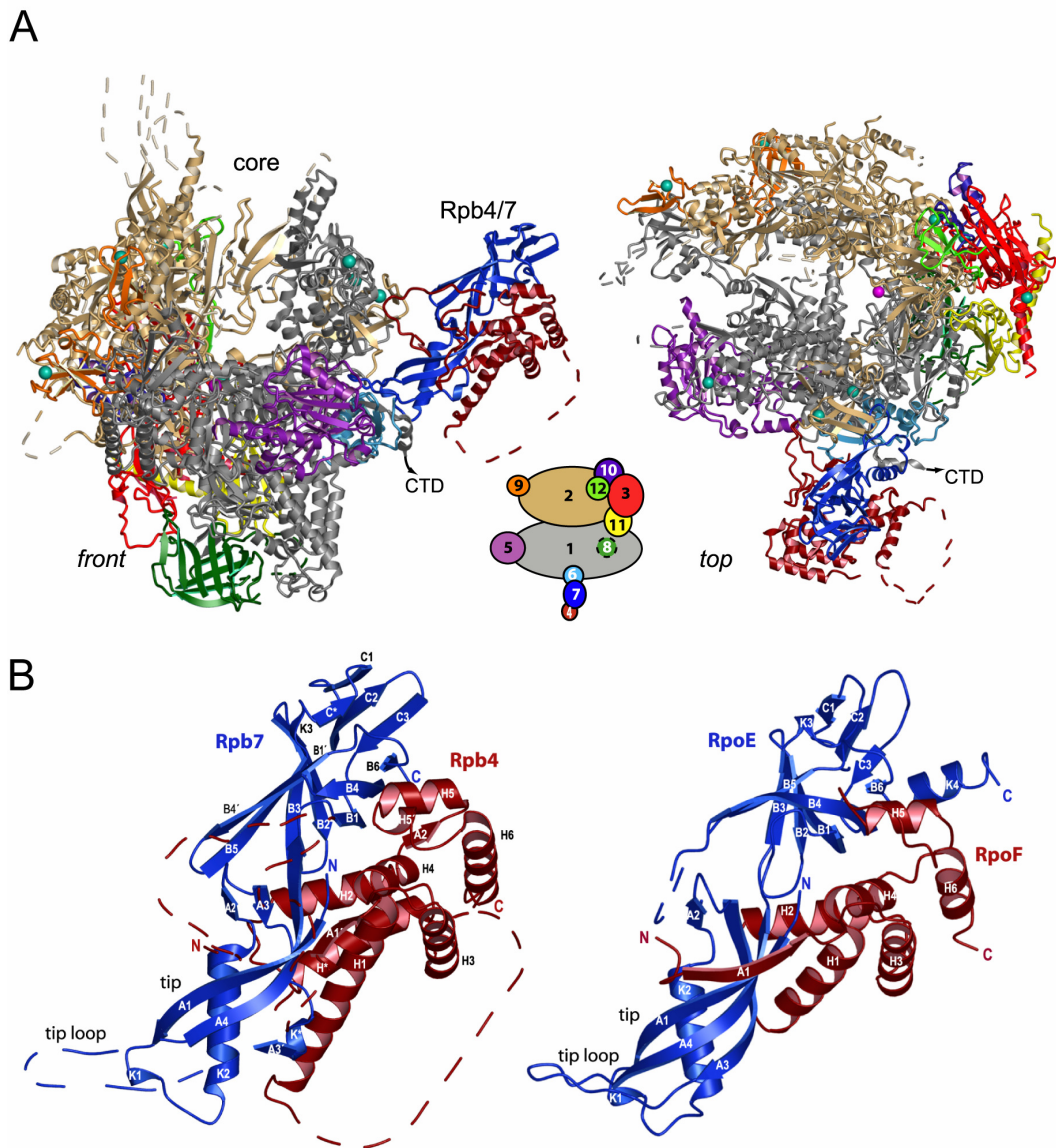
Figure 1.2 Structure of the complete RNA Polymerase II and the Rpb4/7 subcomplex. Adapted from (Armache *et al.*, 2005).

A. Complete RNA polymerase II structure. Ribbon diagram. Two standard views (front and top) are shown. The 12 subunits Rpb1-Rpb12 are colored according to the key below the views. Dashed lines represent disordered loops. Eight zinc ions and the active site magnesium ion are depicted as cyan spheres and a pink sphere, respectively. **B**. Structure of the Rpb4/7 complex. A comparison of the structures of yeast Rpb4/7 (left) (Armache *et al.*, 2005) and archaeal RpoF/E (right) (Todone *et al.*, 2001). Rpb7/RpoE are in blue, and Rpb4/RpoF are in red. Dashed lines represent disordered loops.

1.3. The Rpb4/7 subcomplex

1.3.1. Rpb4/7 structure

The crystal structure of free Rpb4/7 from *S.cerevisiae* at 2.3 Å resolution was published unveiling the detailed architecture of this Pol II subcomplex (Armache *et al.*, 2005). Rpb4/7 showed high similarity to the RpoE/F complex (Figure1.2 b), its counterpart in the archeal RNA polymerase (Todone *et al.*, 2001). More recent X-ray structures of Rpa14/43 and C17/25 (see Chapter II, Figure 3.1), which are Rpb4/7 counterparts in the RNA Pol I and III systems, respectively, confirmed an overall conserved fold of all Rpb4/7-like subunits (Jasiak *et al.*, 2006; Kuhn *et al.*, 2007). The observed structural variations between these subcomplexes may confer specificity to interaction with transcription factors in the different RNA polymerases.

Rpb7 is composed of two domains: the N-terminal tip domain and a C-terminal domain that includes an oligosaccharide-binding (OB) fold and was therefore named the OB domain. The tip domain of Rpb7 harbors a highly conserved "tip loop" localized between an A3' and A4 strand. Together with the helical turn K1, the tip loop contributes to the binding of the subcomplex to the Pol II core. The disordered structure of the loop observed in the free Rpb4/7 becomes ordered upon Pol II binding. Rpb4 binds between the Rpb7 domains and consist also of two domains. The tip associated domain is composed of a single β-strand and three α-helices and is closely connected with the C-terminal HRDC domain. The helicase and RNAse D C-terminal (HRDC) domain adapts very similar orientation in the archeal RpoE/F structure. After showing its different location and the ability to fold independently in vitro in C17, HRDC was acknowledged as a separate structural module (Jasiak *et al.*, 2006). Rpa43 lacks the HRDC domain (Kuhn *et al.*, 2007). Compared with the archeal RpoE and C25 from RNA Pol III, Rpb7 lacks the C-terminal helix and has an additional helix K* inserted into strand A3. With these two features Rpb7 resembles the secondary structure composition of Rpa43, its counterpart in RNA Pol I (Geiger *et al.*, 2008). Rpb4 additionally contains a non-conserved amino-terminal extension, a longer helix H1, and an insertion between helices H1 and H2 (H1-H2 insertion), which comprises a long disordered loop and an additional helix H*. The H* helix, which is not available in the RpoF, is present in C17 and replaced by an about 30 amino acids loop in Rpa14.

1.3.2. The function of Rpb4/7 subcomplex

The core RNA Pol II is sufficient for elongation and promoter-independent initiation of transcription (Ruet *et al.*, 1980; Rosenheck & Choder, 1998) but requires the Rpb4/7 subcomplex for promoter-directed initiation in vitro (Edwards *et al.*, 1991; Orlicky *et al.*, 2001). The Rpb4 and Rpb7 were identified in *S.cerevisiae* as the 4th and 7th biggest subunits of the RNA Pol II complex, respectively (Woychik & Young, 1989). The two subunits form a heterodimer, that can dissociate from the core enzyme (Edwards *et al.*, 1991; Cramer *et al.*, 2008). It is substoichiometric with respect to the Pol II core (Kolodziej *et al.*, 1990). In *S.cerevisiae*, only 20 % of Pol II seems to contain Rpb4/7 in exponentially growing cells. In stationary phase and during stress conditions 12-subunit Pol II becomes the most abundant form of the enzyme (Choder & Young, 1993). In contrast, *S.pombe* and human Rpb4/7 are stoichiometrically associated with Pol II during all growth phases (Sakurai & Ishihama, 1997; Khazak *et al.*, 1998). Whereas Rpb7 is essential for viability of the yeast *Saccharomyces cerevisiae*, Rpb4 is not, but it becomes essential at temperature extremes (Woychik & Young, 1989). At the same time Rpb4 is not required for cell survival during oxidative or osmotic stress (Maillet *et al.*, 1999) or nitrogen starvation (Pillai *et al.*, 2001). The temperature-sensitive phenotype of the ΔRpb4 strain can be partially suppressed by Rpb7 overexpression (Maillet *et al.*, 1999; Tan *et al.*, 2000). Rpb4/7 binds single-stranded nucleic acids and mediates a step during initiation subsequent to promoter DNA binding (Orlicky *et al.*, 2001). These observations are consistent with the idea that Rpb4/7 is present during initiation at promoters but then dissociates from Pol II. However, evidence has accumulated for additional functional roles of Rpb4/7 during transcription (partially reviewed in (Choder, 2004)). Rpb7 remains associated with early elongation complexes (Cojocaru *et al.*, 2008) and binds Nrd1, a protein involved in RNA 3'-end processing (Mitsuzawa *et al.*, 2003). Rpb4/7 is observed to shuttle between the nucleus and cytoplasm (Selitrennik *et al.*, 2006) and is proposed to play roles in P-body function and mRNA decay (Lotan *et al.*, 2005). Recent data show that the Rpb4/7 subcomplex can be crosslinked to the transcribed region, and loss of Rpb4 decreases the association with 3'-processing factors and alters usage of the polyadenylation site at a tested gene (Runner *et al.*, 2008). This suggests that Rpb4/7 is an integral part of the Pol II enzyme and is required not only for initiation, but also for 3'-RNA processing at the end of transcription.

1.4. Aim of this study

Pol II is the enzyme responsible for mRNA synthesis during transcription of protein-coding genes in eukaryotic cells. The structure of the complete 12-subunit Pol II is known and consists of a ten-subunit core enzyme, which includes the active center, and the peripheral heterodimer of subunits Rpb4 and Rpb7 (Rpb4/7 subcomplex) (Cramer *et al.*, 2001; Armache *et al.*, 2003; Bushnell & Kornberg, 2003; Armache *et al.*, 2005). *In vitro*, yeast Rpb4/7 is required for transcription initiation, can dissociate from Pol II, and is dispensable for catalytic RNA elongation (Edwards *et al.*, 1991). These observations are consistent with the idea that Rpb4/7 is present during initiation at promoters but then dissociates from Pol II.

However, evidence has accumulated for additional functional roles of Rpb4/7 during transcription. Rpb7 remains associated with early elongation complexes (Cojocaru *et al.*, 2008) and binds Nrd1, a protein involved in RNA 3'-end processing (Mitsuzawa *et al.*, 2003). Rpb4/7 can be crosslinked to the transcribed region, and loss of Rpb4 decreases the association with 3'-processing factors and alters usage of the polyadenylation site at a tested gene (Runner *et al.*, 2008). Thus, recent data suggest that Rpb4/7 is an integral part of the Pol II enzyme and is required not only for initiation, but also for 3'-RNA processing at the end of transcription.

To investigate whether Rpb4/7 generally associates with Pol II *in vivo*, the chromatin immunoprecipitation experiments in yeast was carried out and coupled to tiling microarray analysis at a technical resolution of 32 base pairs (ChIP-chip). The occupancy profiles of Rpb7 and the Pol II core subunit Rpb3 were obtained and compared to show whether complete Pol II, including Rpb4/7, associates with DNA genome-wide.

2. Results and discussion

2.1. Temperature-sensitivity tests of the yeast strains

All the ChIP experiments were performed on *Saccharomyces cerevisiae* W303 and S288C strains carrying a tandem affinity purification (TAP) tag sequence fused to the C-terminus of the essential genes encoding the Rpb3 or Rpb7 subunit of RNA Pol II (Chapter III, 3.1). The temperature-sensitivity tests were performed to investigate whether the introduced TAP-tag influences transcription and yeast viability by disturbing the integrity of the Pol II complex or its interactions with other proteins. Growth of each strain at three different temperatures (24ºC, 30ºC and 37ºC) was tested (Chapter III, 3.1.1) and compared to that of the W303 wild type strain. Both, the wild type and TAP-tagged strains showed similar growth therefore demonstrating no influence of the introduced tag on the functionality of the transcription machinery (Figure 2.1).
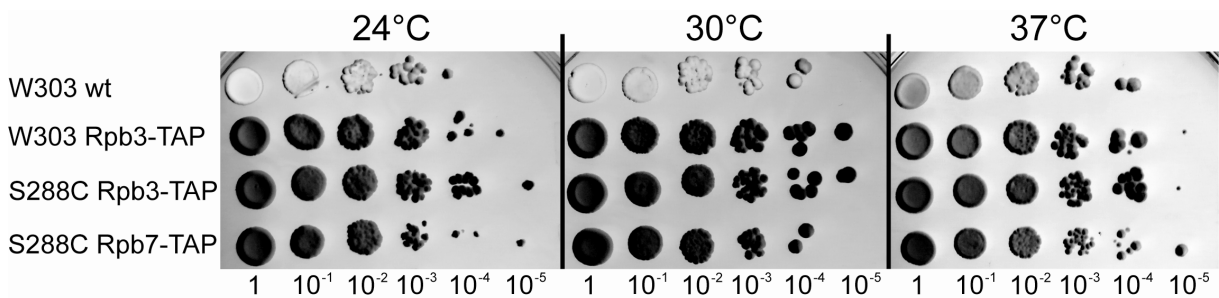
Figure 2.1 Temperature-sensitivity test of the TAP-tagged W303 and S288C yeast strains.
Serial dilutions of W303 wild type (wt), W303 Rpb3-TAP, S288C Rpb3-TAP and S288C Rpb7-TAP strain grown on YPD plates at 24ºC (left), 30ºC (center) and 37ºC (right).

2.2. Monoclonal antibody selection

2.2.1. Western blot analysis

Monoclonal antibodies against Rpb4/7 were produced by Dr. Elisabeth Kremmer from Helmholtz Zentrum München, Institut für Molekulare Immunologie, Hämatologikum (Chapter III, 3.2). The specificity of the 14 different hybridoma supernatants containing antibodies against Rpb4/7 was tested (for details see Chapter III, 3.3). Only antibodies number 1, 2, 3, 4, 8, 9, 10, 13 and 14 could recognize yeast proteins from the W303 wt yeast strain extract and were further analyzed. Three antibodies, number 1, 2 and 9, showed high specificity in Western blot analysis. They recognized both, recombinant and yeast endogenous Rpb4/7 (Pol

II-TFIIF) and did not interact with the Rpb4/7 homologues in Pol I and Pol III (Figure 2.2).
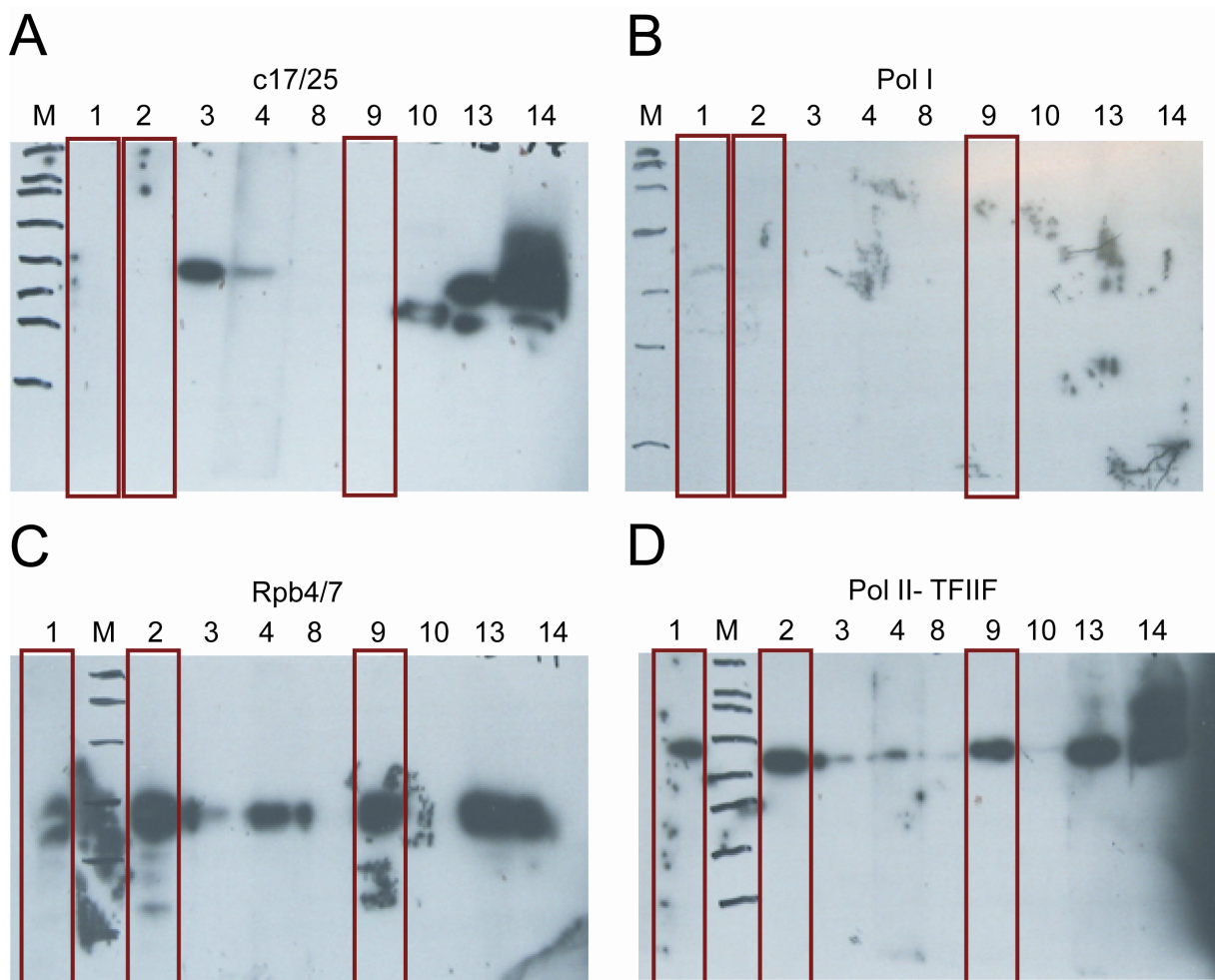
Figure 2.2 Wester blot analysis of the monoclonal antibodies specifity.

The specificity of the monoclonal antibodies against Rbp4/7 was investigated by testing their interaction with C17/25 (**A**) of RNA Pol III, RNA Pol I (**B**), recombinant Rpb4/7 (**C**) and a yeast endogenous Rpb4/7 from RNA Pol II – TFIIF preparation (**D**). The numbers of the antibodies are indicated above the gels. Lanes with the specific antibodies 1, 2 and 9 are indicated with the red boxes.

2.2.2. Chromatin immunoprecipitation tests

In contrast to immunoprecipitation, where an antibody binds the folded and cross-linked protein complexes, proteins used for a Western blot analysis are unfolded during SDS-PAGE. To test if the specific antibodies recognize also folded Rpb4/7 and whether the recognized epitope is accessible also in the context of the whole Pol II, chromatin immunoprecipitation (ChIP) experiments were performed (Chapter III, 3.4). To ensure, that there is no unspecific binding of the chromatin to the beads, mock IP experiments were performed. Non-template IP showed that there are no unspecific DNA impurities coming from buffers and reagents. The DNA obtained in the ChIP experiments was analysed with a standard PCR (Chapter III, 3.4.8). DNA

was diluted in ready PCR sample to 1:250 for Input samples, two dilutions, 1:250 and 1:500 were used for IP, mock and non-template IP samples. Two loci, the middle of Adh1 gene on chromosome XV and Reference region (Ref) on chromosome V were analysed. PCR analysis of the IP samples should give a signal only on the highly transcribed Adh1 gene, whereas both, Adh1 and non-transcribed Reference region should be present in the Input samples.

Only antibody (Ab) 2 and 9 gave a positive signal with the Rpb4/7 complex and the Adh1 gene (Figure 2.3, lanes 2-5). No signal for the Ref region could be observed indicating a high specificity of the antibody binding. The lack of PCR products in the mock and non-template IP samples (Figure 2.3, Lanes 6-10) shows that there is no unspecific DNA bound to the beads or coming from the reagents used for the experiments.
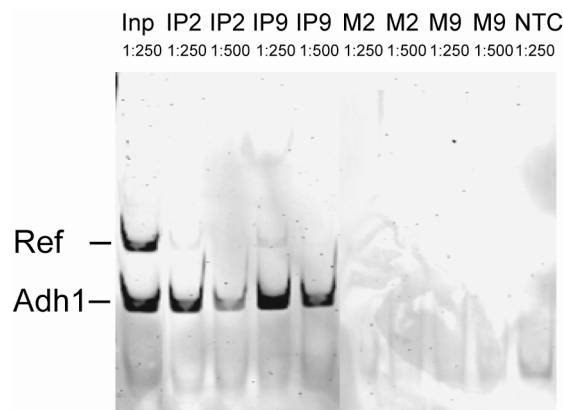
Figure 2.3 PCR analysis of the ChIP experiments with the monoclonal antibodies against Rpb4/7.
PCR products were analysed by gel electrophoresis (Chapter III, 3.4.8). Sample names and dilutions are indicated above the lanes. Inp – Input sample; IP2 – IP sample of ChIP with Ab 2; IP9 - IP sample of ChIP with Ab 9; M2 – Mock IP sample of ChIP with Ab 2; M9 – Mock IP sample of ChIP with Ab 2; NTC – non-template control.

2.3. Genome-wide ChIP of Pol II subunits

To investigate whether the complete Pol II is the form of the enzyme that associates with the yeast genome, or whether the complete Pol II and the core Pol II associate with different regions of the genome, the ChIP coupled to occupancy profiling with high-resolution tiling microarrays (ChIP-chip) was used. Yeast strains used for the experiments contained TAP tags at the C-termini of the integral core subunit Rpb3 and the essential Rpb7 subunit, which is part of the Rpb4/7 subcomplex. Both subunit C-termini are exposed in the Pol II structure (Armache *et al.*, 2005), suggesting that the tag would not interfere with Pol II function *in vivo*. Additionally the growth of all the yeast strains was tested in the temperature-sensitivity test (Chapter III, 2.1). No negative influence of the introduced tag on the yeast viability could be observed.

Obtained microarray data have been submitted in MIAME-compliant form to the Gene Expression Omnibus database (http://www.ncbi.nlm.nih.gov/projects/geo) under accession number GSE12060.

2.4. Analysis and quality of ChIP-chip data

To investigate the reproducibility of the ChIP-chip data, three biological replicate experiments with the S288C Rpb3-TAP strain were carried out. The reproducibility of the obtained data was very high, resulting in Pearson correlations between individual profiles of around 0.9 (Table 2.1). The Pearson coefficient is 1 for a perfect correlation and 0 for uncorrelated signals. Similarly, two biological replicates of S288C Rpb7-TAP profiles were obtained and show very good agreement (Pearson correlation 0.90, Table 2.1). Noise could be dramatically reduced using dual color microarray technology and replica measurements with interchanged dyes (Yang & Speed, 2002) (Figure 2.4 a-c; Table 2.1).

Table 2.1 Pearson correlation coefficients between occupancy profiles

Comparison	Profile 1 (subunit, tag, strain)	Profile 2 (subunit, tag, strain)	Pearson correlation, unsmoothed	Pearson correlation, smoothed
Biological replicates[1]	Rpb3-TAP, S288C	Rpb3 TAP, S288C	0.88	0.92
Rpb3 vs. Rpb7[2]	Rpb7-TAP, S288C	Rpb3-TAP, S288C	0.91	0.93
Different strains[1]	Rpb3-TAP, S288C	Rpb3-TAP, W303	0.90	0.92
TAP tag vs. antibody[1]	Rpb7-TAP, S288C	Rpb4/7 AB, W303	0.72	0.79

[1] Profiles are single measurements with the ChIP DNA tagged with the Cy5 dye.
[2] Profiles for Rpb3 and Rpb7 are averaged over three and two measurements, respectively, one measurement in each case taken with interchanged fluorescent dyes. The averaging over replicas as well as the smoothing reduces noise and leads to higher correlation coefficients.
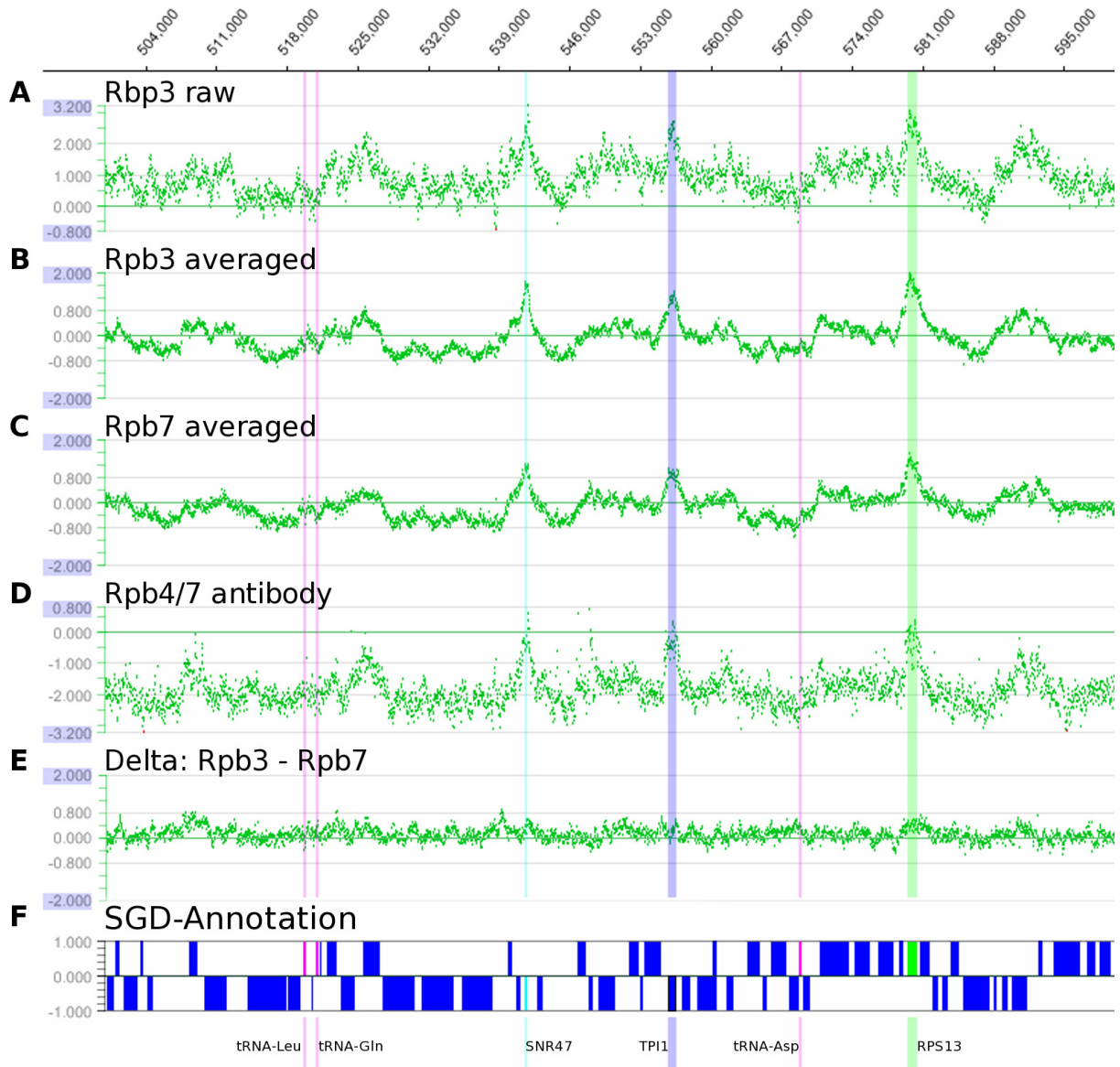
Figure 2.4 Genome-wide occupancy profiling of Pol II subunits.

Profiles were obtained by chromatin immunoprecipitation coupled to high-resolution tiling microarray analysis. A representative 100 kbp sample on chromosome 4 (genomic positions 500,000-600,000) of the profile for the Rpb3 and Rpb7 subunits is depicted. Each green dot represents the signal for a single oligonucleotide probe on the tiling array, which has one probe every 32 bps. (**A**) Raw data trace for the TAP-tagged Rpb3 core subunit of Pol II. The signal intensity is the logarithm (base 2) of the fluorescence signal of Cy5-labeled Rpb3-bound DNA divided by the signal for Cy3-labeled genomic background DNA. (**B**) Average over three biological replicate traces for Rpb3, one of which with interchanged fluorescent dyes. (**C**) Average over two biological replicate traces for Rpb7, one with interchanged dyes. (**D**) Pol II occupancy profile generated with a monoclonal antibody against Rpb4/7. (**E**) The difference signal between averaged Rpb3 and Rpb7 occupancy profiles (**B**, **C**) fluctuates around a near-zero baseline. (**F**) Annotation of genomic features according to the Saccharomyces Genome Database. Genes encoding the intron-containing ribosomal protein RPS13 (green), the snoRNA SNR47 (cyan) and the housekeeping gene triose phosphate isomerase TPI1 show high Rpb3 occupancy at our growth conditions. Other protein-coding genes (blue) show lower Pol II occupancy. Three different tRNA genes (magenta) are transcribed by Pol III and show no increased Pol II occupancy.

2.5. Occupancy profiles for Rpb3 and Rpb7 are essentially identical

The replica-averaged, noise-reduced profiles for Rpb3 and Rpb7 were compared. Differences in the two profiles would indicate that part of the genome associates with the complete Pol II, and part of the genome associates with the core enzyme, and thus Rpb4/7 is dissociable *in vivo*. In contrast, a high similarity of the two profiles would indicate that Rpb4/7 is always part of the DNA-associated Pol II *in vivo*.

Comparison of the average, noise-reduced occupancy profiles for Rpb3-TAP and Rpb7-TAP showed that they are highly correlated throughout the genome (Figure 2.4 b,c). There are only small differences between the profiles (Figure 2.4 e, Figure 2.5). Weak systematic deviations from zero occur only for a small fraction of genomic locations with highest signal intensity (Figure 2.5, see discussion below). A Pearson correlation coefficient of 0.91 between the averaged Rpb3 and Rpb7 profiles compares favorably with the correlation obtained for biological replicates of the same profile (0.88, Table 2.1). Thus, statistically the agreement between the two average profiles of Rpb3 and Rpb7 was as good as the agreement between individual replicate profiles of the same strain. The near identity of the two profiles, within experimental errors, indicates that Rpb4/7 is present in DNA-associated Pol II genome-wide.
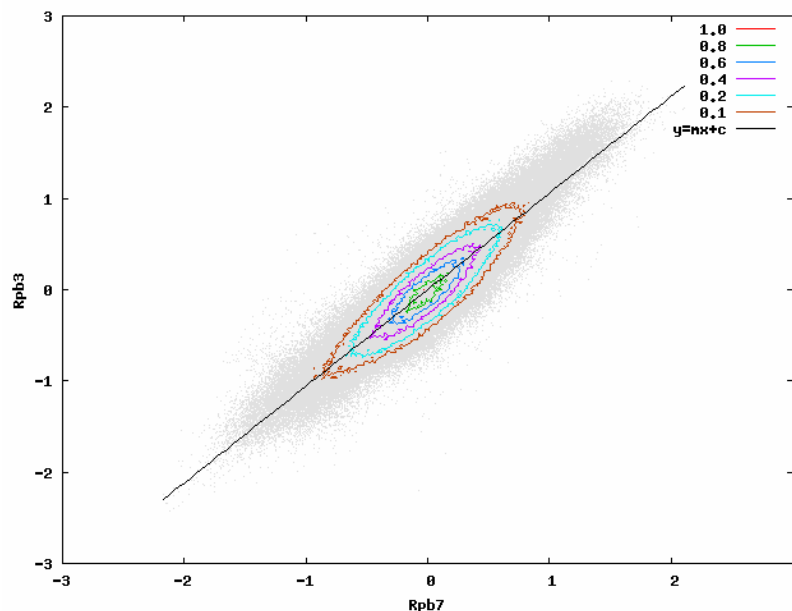
Figure 2.5 Rpb3 and Rpb7 occupancies are highly correlated.

The figure shows the Rpb3 versus Rpb7 signals (calculated as log base 2 of ChIP signal divided by genomic background signal). Each tiling array probe contributes a grey dot. Contour lines go from 0.8 to 0.1 of maximum density. Approximately 90% of probes are contained within the 0.1 contour. The correlation is very high (Pearson correlation coefficient 0.91), comparable to the correlation of Rpb7 signal between biological replicates and different strains. The slope of the regression line is 1.06. This factor is used to scale up the Rpb7 signal when calculating the difference signal (Rpb3 – Rpb7) in Figures 2.4 e and 2.6 c.

2.6. The profiles are not systematically influenced by the type of yeast strain

The obtained occupancy profiles were investigated whether thay change when a different strain of yeast is used. The Rpb3-TAP profile obtained with yeast strain S288C was compared to a profile obtained under identical conditions with a W303 strain. The obtained profiles were highly similar, resulting in a Pearson correlation of 0.90 (Table 2.1), showing that the type of yeast strain did not influence the results.

2.7. The profiles are not influenced by the affinity tag

It is possible that the cellular function of Rpb7 is perturbed by its fusion to the TAP tag. To investigate this, the recombinant stoichiometric Rpb4/7 subcomplex was purified as described (Armache *et al.*, 2003) and used for raising of the monoclonal antibodies against the pure subcomplex (Chapter III, Experimental procedures). Obtained monoclonal antibody was used to measure an unbiased occupancy profile for Rpb4/7. The resulting occupancy profile contained more noise than the profiles obtained from TAP-tagged strains, but it did not show any systematic disagreement with the profile obtained from TAP-tagged strains (Figure 2.4 d) resulting in a Pearson correlation of 0.72 (0.79 for the smoothed traces, Table 2.1). Thus, the profiles obtained from TAP-tagged strains reflect those of unperturbed yeast cells, and were not considerably influenced by the presence of the TAP tag.

2.8. Correlation of the Pol II occupancy profiles with genome features

The ChIP-chip data contain a lot of information beyond the observation that Rpb4/7 is a component of DNA-associated Pol II *in vivo*. Analyses of Pol II occupancy profiles over a part of the yeast genome or at lower resolution were presented recently (Liu *et al.*, 2005; Radonjic *et al.*, 2005; Steinmetz *et al.*, 2006), but a comparison of these published profiles with our data is hampered due to the use of different experimental conditions and technological platforms.

To analyze which genomic features correlate with Rpb3 occupancy, all genomic regions were identified with a ChIP-chip signal above an occupancy threshold that was varied continuously (Figure 2.6 a). We added up the base pairs within these regions belonging to the genomic features annotated in the Saccharomyces Genome Database (version 11 March 2008). The colored areas in Figure 2.6 a indicate the fractions of the regions annotated by various genomic features at a chosen Pol II threshold. A threshold of zero corresponds to equal fluorescent signals from Rpb3 and the genomic background. The solid line shows the proportion of the genome

above the threshold. Figure 2.6 b shows how much certain genomic features are overrepresented in regions with Pol II occupancy above a given threshold.
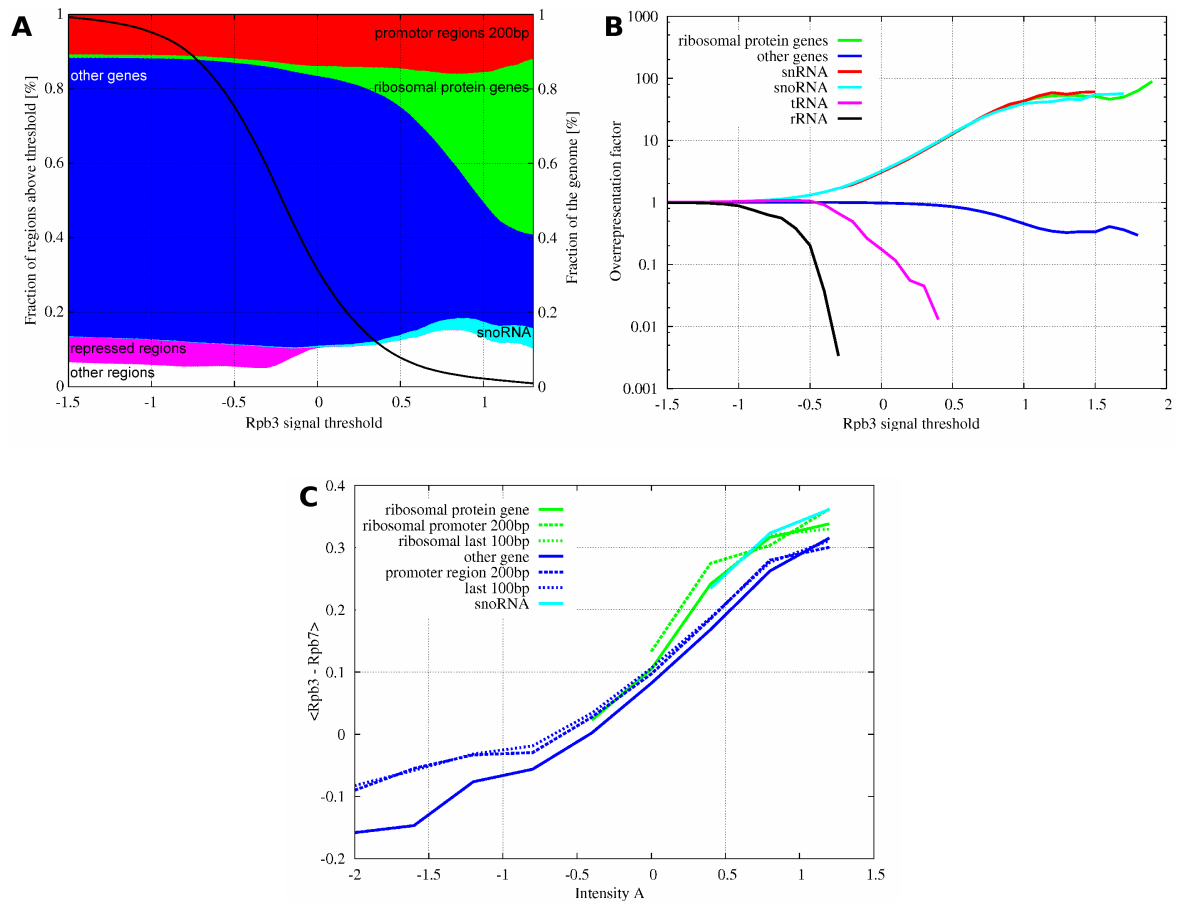
Figure 2.6 Distribution of genomic regions above a Pol II (Rpb3-TAP) occupancy threshold.

A. The thick black trace shows the fraction of the genome within regions above the Rpb3 occupancy threshold. For example, ~99% of the whole genome have occupancy signals above a threshold of -1.0, ~30% have signals above 0 (this threshold corresponds to equal fluorescent signals from Rpb3 and the genomic background). At each threshold, the height of the colored areas indicates the fraction of above-threshold regions annotated with the given genomic feature. Non-ribosomal protein genes (blue) make up almost 80% of the genome but they constitute only 25% in regions with a Rpb3 occupancy above 1.0. Ribosomal protein genes (green) occupy only a small fraction of the genome, yet make up 35% of regions with an Rpb3 occupancy above 1.0. The fraction of all genes (ribosomal and non-ribosomal) is approximately constant over a large range of thresholds, as is the fraction in promoter regions, defined as the 200 bps upstream of genes (red). The most overrepresented features in highly occupied regions are snoRNAs (cyan) and ribosomal genes (green). Repressed regions (magenta) summarize LTR retrotransposons, transposable elements and telomeres and are absent in regions above an occupancy threshold of 0. **B**. Overrepresentation ("enhancement factor") of selected genomic features in regions above the occupancy threshold. Colors are as in A except for the snRNAs (red). "Other regions" encompass intergenic regions of unknown function and more rarely occurring genomic features. **C**. Dependence of the difference signal (Rpb3 – Rpb7) on the average signal intensity A = (Rpb3+Rpb)/2. Colors are as in A and B, solid traces refer to the coding region, dashed traces to promoter regions (i.e. 200 bp upstream of coding regions), dotted traces to the last 100 bp of coding regions. Since the signals are calculated as logarithms (base 2) of the ChIP fluorescence divided by the genomic background fluorescence, a difference signal of 0.3 would correspond to a ratio of Rpb7 to Rpb3 occupancy of 2-0.30 = 0.81

2.9. Pol II distribution over the genome

This analysis provides general insights into the global distribution of Pol II in growing yeast cells. First, most Pol II is bound to highly transcribed genes that encode snRNAs, snoRNAs, or ribosomal proteins (Figure 2.6). Second, genomic repeat regions that are expected to be transcriptionally silenced show very low Pol II occupancy. Third, Pol III-transcribed genes that encode tRNAs do not show Pol II enrichment. Fourth, the fraction of all genes and promoter regions is approximately constant over a large range of Pol II occupancy thresholds (Figure 2.6 a). This supports previous reports, which have found no evidence for widespread accumulation of Pol II by pausing at promoters in yeast, in contrast to higher eukaryotes (Wade & Struhl, 2008). However, our findings are also consistent with other ChIP-chip studies showing the preferential location of Pol II on promoter-bearing intragenic regions in yeast cells in stationary phase (Radonjic *et al.*, 2005) as our data were obtained from yeast cells harvested in their logarithmic growth phase.

2.10.   Persistent presence of Rpb4/7

While the results of this study were submitted to the publishing, a study comparing the genome-wide distribution of core Pol II and Rpb4/7 subcomplex became available (Verma-Gaur *et al.*, 2008). This study showed at lower resolution that Rpb3 and Rpb4 profiles are similar and generally agrees with our findings. The authors also note that Rpb4 can be underrepresented with respect to Rpb3, an effect that is more pronounced for shorter genes. However, it could not be confirmed that this is a significant or systematic effect, as demonstrated below.

The ratio of Rpb7 to Rpb3 does not fall below 0.78, and such deviations from 1 are limited to a very small fraction of the genome (Figure 2.6 c). These minor local deviations were investigated whether they show systematic features, and thus may be biologically significant. The Rpb7-Rpb3 difference signal was analyzed separately for several genomic categories, including ribosomal protein genes, nonribosomal protein genes, and snoRNA (Figure 2.6 c). This revealed that the difference signal increases with increasing occupancy on the x-axis. The higher occupancy of Rpb3 versus Rpb7 for ribosomal proteins can be explained by the generally higher total Pol II occupancy on ribosomal protein genes. Rpb7 occupancy also did not depend on the position along the gene. The promoter regions and 5' ends of genes show essentially the same difference signal as the coding regions. Thus the ratio of Rpb3 to Rpb7 occupancy neither depends on the genomic feature type nor on the position along genes. Thus there is no evidence for a functional significance of the minor local deviations between Rpb3 and Rpb7 levels.

2.11. Functional roles of Rpb4/7

The data presented in this work suggest that Rpb4/7 is generally required throughout the transcription cycle *in vivo*, and argue against models that propose the dissociation of Rpb4/7 during transcription. Obtained results are consistent with functional roles of Rpb4/7 during initiation (Edwards *et al.*, 1991; Orlicky *et al.*, 2001) as well as 3'-RNA processing (Runner *et al.*, 2008). They are also consistent with microarray-based gene expression analysis of yeast strains lacking the gene for Rpb4, which suggested that Rpb4 is globally required for Pol II transcription (Miyao *et al.*, 2001; Pillai *et al.*, 2003). They are further consistent with the observation that Rpb4 is essential in the fission yeast *S. pombe*, and that Rpb4/7 is present in *S. pombe* Pol II in stoichiometric amounts during exponential cell growth (Sakurai *et al.*, 1999).

Finally, the results presented here are not inconsistent with the recent observations that Rpb4/7 shuttles between the nucleus and cytoplasm (Selitrennik *et al.*, 2006), and that Rpb4 and Rpb7 play roles in P-body function and mRNA decay (Lotan *et al.*, 2005), because it is possible that Rpb4/7 dissociates from core Pol II during transcription termination, and that there is a cellular pool of free Rpb4/7. It also remains possible that Rpb4/7 dissociates under certain growth conditions.

3.   Experimental procedures

3.1. Yeast strains

Experiments were performed with Saccharomyces cerevisiae strains W303 wild type and W303 carrying a tandem affinity purification (TAP) tag sequence fused to the C-terminus of the gene encoding Rpb3. In addition, *S. cerevisiae* S288C strains RPB3-TAP and RPB7-TAP, carrying a TAP-tag sequence fused to the C-terminus of the genes encoding Rpb3 and Rpb7 respectively were used. Table 3.1 summarises detailed list of yeast strains and their genetic background.

Table 3.1 Yeast strains

strain	description	source
W303 wt	MATa, ura3-1, trp1-1, his3-11,15, leu2-3,112, ade2-1, can1-100, GAL+	Open Biosystems
W303 Rpb3-TAP	MATa, ura3-1, his3-11,15, leu2-3,112, ade2-1, can1-100, GAL+, RPB3-TAP::TRP1	Open Biosystems
S288C Rpb3-TAP	ATCC 201388:MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0	Open Biosystems
S288C Rpb7-TAP	ATCC 201388:MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0	Open Biosystems

3.1.1. Temperature-sensitivity tests

Single colony of yeast strain was picked from a fresh plate and resuspended in 500 µl sterile water by gentle pipetting. Seven sequential 1 to 10 dilutions of the resuspended colony in sterile water were prepared and 5 µl of each dilution was spotted in the row on three YPD plates. After drying under the hood for 5 minutes, plates were placed in 24ºC, 30ºC and 37ºC incubators, respectively, and were grown for next three days.

3.2. Production of monoclonal antibodies against Rpb4/7

Monoclonal antibodies against Rpb4/7 were produced by Dr. Elisabeth Kremmer form Helmholtz Zentrum München, Institut für Molekulare Immunologie, Hämatologikum as a part of service.

Lou/C rats were immunized subcutaneously and intraperitoneally with a mixture of 50 µg of purified recombinant Rpb4/7-His fusion protein (Armache *et al.*, 2003), 5 nmol CpG oligonucleotide (ODN 2006, TIB Molbiol, Berlin, Germany), 500 µl PBS and 500

µl IFA. After a six-week interval a final boost without adjuvant was given three days before fusion of the rat spleen cells with the murine myeloma cell line P3X63-Ag8-653. Hybridoma supernatants were tested in an ELISA using bacterial extracts from *E. coli* expressing either the Rpb4/7 fusion protein or a His-tagged fusion protein.

3.3. Monoclonal antibody selection

3.3.1. Yeast extract preparation

A starter culture of W303 wt strain was grown from single colony in 50 ml YPD medium over-night in 30ºC. Fresh pre-warmed YPD medium was inoculated with a starter culture to OD 0.2 and shaken (180 rpm) at 30 °C until the culture reached OD 1.5. 10 OD of cells were harvested by centrifugation 5 min, 4000 rpm, 4ºC. After washing the pellet with ice-cold water the cells were resuspended in 200 µl of hot (80ºC) SDS loading buffer by vortexing for 30 sek. Cell lysis was performed using 100 µl of glass beads by several sequential vortexing (60 sek) and boiling (2 min in 95ºC) steps, followed by centrifugation 5 min, 13 000 rpm, 4ºC in order to get rid of cell debris. 5 µl per well of cell lysate was used for the SDS-PAGE gels.

3.3.2. Protein samples preparation

Rpb4/7, 12 subunits RNA Polymerse II/TFIIF complex and RNA Polymerse I used for specificity tests of the antibodies were provided by Florian Brückner, Anass Jawhari and Claus-Dieter Kuhn, respectively, and were purified as described before in (Armache *et al.*, 2003) and (Kuhn *et al.*, 2007). Rpc17/25 subcomplex was purified according to the described protocol (see Chapter II, Experimental procedures).

3.3.3. Protein separation by SDS-PAGE

To separate distinct protein bands SDS-PAGE with 15% acrylamide gels (Laemmli, 1970) was performed. Before loading onto the gel proteins were unfolded by adding β-mercaptoethanol to the SDS loading dye. After electrophoresis gels were further subjected to blotting procedures.

3.3.4. Western blot analysis

After electrophorsis protein gels were blotted onto ethanol pre wet PVDF membranes (Roth) for 1 h at 90 V in 4ºC, using a transfer chamber (BioRad Trans-Blot Cell). In order to indentify position of single lanes, blotted membrane was shortly stained with Ponceu S solution and cut to one lane broad stripes. After several washing steps with water resulting removing of the dye, the membrane was blocked with Blocking

Solution, three times for 20 min in room temperature, followed by over-night incubation with Primary Antibody Solution in 4ºC (each membrane stripe incubated separately with different antibody). Antibody coated membranes were washed three times 15 min in 1 x PBS solution and incubated for 2 h in room temperature with Secondary Antibody Solution followed by second PBS washing round (3 x 15 min). Detection was performed with the ECL Plus reagents (Amersham Pharmacia Biotech) following the manufacturers instructions.

Transfer buffer 2l

24 g	Glycine
8 g	Tris HCl
400 ml	Ethanol
~ 1600 ml	H_2O

Blocking solution

2 % (w/v)	skimmed milk powder in 1 x PBS

Primary Antibody Solution

25 x diluted	Hybridoma supernatant solution
2 % (w/v)	skimmed milk powder in 1 x PBS

Secondary Antibody Solution

4500 x diluted	Goat anti Rat IgG - IgM coupled to HRP (horse-radish-peroxidase)
2 % (w/v)	skimmed milk powder in 1 x PBS

Ponceu S

0.1% (w/v)	Ponceau S in 5% (v/v) acetic acid

1 x PBS 1l

8 g	NaCl
0.2 g	KCl
1.44 g	Na_2HPO_4
0.24 g	KH_2PO_4
pH 7.4	Adjusted with HCl

3.4. Chromatin immunoprecipiation experiments - ChIP

3.4.1. Media and buffers

All solutions were prepared using double-distilled water. All media and buffers were autoclaved or sterile filtered for storing and usage.

YPD Medium

1.5 %	yeast extract
2 %	peptone
2 %	glucose

FA lysis buffer with and without 2 mM PMSF or 0.5 M NaCl

50 mM	HEPES: adjust pH to 7.5 with KOH
150 mM or 0.5M	NaCl
1 mM	EDTA
1%	Triton X-100
0.1%	sodium deoxycholate
0.1%	SDS

For FA lysis buffer with 2 mM PMSF add 100 mM phenylmethylsulfonyl fluoride (PMSF) in ethanol (store up to 1 year at –80ºC) to a final concentration of 2 mM just before use. For FA lysis buffer with 0.5 M NaCl, change the amount of NaCl added to 0.5 M.

ChIP wash buffer

10 mM	Tris·Cl pH 8.0
0.25 M	LiCl
1 mM	EDTA
0.5%	Nonidet P-40
0.5%	sodium deoxycholate

ChIP elution buffer

50 mM	Tris·Cl pH 7.5
10 mM	EDTA
1%	SDS

TBS 10 x

0.5 M	Tris·Cl pH 7.5
1.5	NaCl

TE buffer

10 mM	Tris·Cl pH 7.5
1 mM	EDTA

3.4.2. Yeast culture

A fresh colony was picked from the plate, and a starter culture was grown over-night in 100 ml YPD medium in 30ºC. Fresh pre-warmed YPD medium was inoculated with a starter culture to OD 0.2 and shaken (180 rpm) at 30 °C until the culture reached log phase (OD 0.7 - 0.8). Protein cross-linking was performed by adding 37% formaldehyde directly into the culture to the final concentration of 1% and slowly shaking for 15 min in room temperature. After stopping the cross-linking by adding 3M Glycine solution, in the final concentration of 2.5%, and incubating the culture for further 30 min in room temperature cells were harvested by centrifugation 5 min, 4000rpm, 4°C and washed three times with ice-cold 1x TBS and once with FA lysis buffer with 2mM PMSF. Cell pellets were flash-frozen in liquid Nitrogen and stored in -80°C up to several months.

3.4.3. Beads preparation

For immunoprecipitation of TAP-tagged Rpb3 and Rpb7, IgG SepharoseTM 6 Fast Flow (GE Healthcare) was used. Before usage the beads were washed three times with ice-cold TBS buffer and one time with FA-lysis buffer.

For precipitation of Rpb4 and Rpb7 from W303 wild type cells, a mixture of Protein A and Protein G Sepharose was incubated with rabbit-anti-rat IgG antibodies for 1 h at room temperature and then incubated with rat monoclonal antibodies against Rbp4/7 for an additional 4-5 h at room temperature. Before adding the chromatin solution the beads were washed with FA lysis buffer.

3.4.4. ChIP experiments

ChIP was performed essentially as described (Aparicio *et al.*, 2005). Briefly, chromatin was sheared with a BioruptorTM UCD-200 (Diagenode) using 25 x 30 sec cycles with 30 sec breaks at an output of 200 W. Before proceeding with immunoprecipitation experiments 20 µl of chromatin solution was put aside and

marked as an Input sample. For immunoprecipitation of TAP-tagged Rpb3 and Rpb7, IgG SepharoseTM 6 Fast Flow (GE Healthcare) was used. Precipitation was performed for 3 h at room temperature. For precipitation of Rpb4 and Rpb7 from W303 wild type cells, antibody-coupled beads (see Chapter III, 3.4.3) were incubated overnight at 4°C with chromatin solution. After immunoprecipitation the beads were washed with FA lysis buffer, FA lysis buffer with 0.5 M NaCl, ChIP wash buffer and TE buffer followed by elution in ChIP elution buffer in 65ºC for 20 min. Eluted IP and Input samples were incubated with Pronase in 42ºC for 3 hours followed by 65ºC for 9 hours for revershal of the crosslinking and protein digestion. Nucleic acids were purified using QIAquick PCR Purification Kit (Qiagen) complemented by PB buffer (Cat. No. 19066, Qiagen) replacing a standard DNA binding buffer with pH indicator provided with the kit. After RNA digestion with RNase A for 30 min in 37ºC DNA was once more purified and concentrated in a Speed Vac to the volume of 10 μl.

3.4.5. Phenol-Chloroform extraction of nucleic acids

In some cases phenol-chloroform extraction of the nucleic acids was applied as an alternative for the QIAquick PCR Purification Kit. Samples were transferred into Maxtract High Density 1.5 ml tubes (Qiagen) and after adding 1 volume of phenol-chloroform solution (Roth) and intensive vortexing, were centrifuge 13 000 rpm/10 min/RT. The upper, water phase was than transferred into a fresh tube with 1/10 of the sample volume of 3M Na Acetate and filled up with 5 volumes of ice cold Ethanol (100%). After mixing and 20 min incubation on the dry ice, precipitated DNA was pelleted by centrifugation 10 min/13 000 rpm/RT. DNA pellet was washed with 80% ice cold Ethanol and dried for 20 min under the hood. Dried DNA was resuspended in appropriate volume of water or TE buffer and stored in -20ºC upon usage.

3.4.6. ChIP quality control experiments

In order to test the specificity of the monoclonal antibody and check for any unspecific binding of the chromatin to the beads, mock IP experiments were performed. For W303 and S288C strains, a mixture of Protein A and Protein G Sepharose coated with rabbit-anti-rat IgG antibodies or a mixture of Protein A and Protein G Sepharose alone, respectively, was used. To ensure, that there are no unspecific DNA impurities coming from buffers and reagents used during experiments non-template IP experiments, during which the beads were incubated with the FA lysis buffer instead of the chromatin solution, were performed. All the control experiments were prepared parallel to the ChIP experiments, according to the protocol described above and were further analysed with standard or quantitive PCR. DNA from a single W303 wt mock IP experiment was amplified and analysed on a microarray according to the protocols used for IP samples and described below.

3.4.7. DNA amplification

DNA samples were amplified and re-amplified with GenomePlex® Complete Whole Genome Amplification (WGA) Kit and GenomePlex® WGA Reamplification Kit (Sigma) using the Farnham Lab WGA Protocol for ChIP-chip (http://www.genomecenter.ucdavis.edu/farnham /pdf/8-18-06WGA.pdf). DNA quantity and quality control was performed with a ND-1000 Spectrophotometer (NanoDrop Technologies). For checking the DNA fragments size distribution 800 ng of reamplified sample was loaded onto 1% Agarose gel. After gel electrophoresis Ethidium Bromide stained-DNA was was observed under the UV light.

3.4.8. Standard pcr analysis

PCR Master Mix (1x)

0.05 µl	forward primer
0.05 µl	reverse primer
1 µl	10 x Polymerase Buffer
1 µl	10 mM $MgCl_2$
1 µl	2 mM dNTP mix
0.15 µl	Taq Polymerase
1.75 µl	H_2O

Table 3.2 Oligonucleotides used for PCR

name	5'→3' sequence
Adh1-M_Forward	CTTGATGGCCGGTCACTGGGTTGC
Adh1-M_Reverse	ATGAAGACTTCACCACCGATGGATC
Ref1_Forward	TGTACTCTCCCACCATTGGGTATTA
Ref1_Reverse	AGTGGTTTTTATTCGAAAGTTGTGGA

Table 3.3 PCR reaction detalis

Cycle No.	step	temperature	duration
	1	94	15 min
26	2	94	30 sek
		55	30 sek
		72	30 sek
	3	72	10 min
	4	4	Pouse

PCR samples were prepared by mixing 5 µl of master mix with 5 µl of DNA sample in 250 µl thin walled PCR tubes. The end dilution of DNA in ready PCR sample was 1:250 or 1:250 and 1:500 for Input samples or IP, mock and non-template IP samples, respectively. To test the purity of the reagent solution non-template PCR control, with 5 µl of sterile water replacing DNA template, was performed. Each PCR sample was independently tested by two pair of primers listed in Table 3.2. After PCR amplification (for the details see Table 3.3) corresponding PCR samples of each primer pair were joined together and mixed with the 5 x DNA Loading Dye (Fermentas). 5 µl of each sample was loaded onto Novex® 4-20 % gradient 1 x TBE gels (Invitrogen). Gel electrophoresis was performed following the manufacturers instructions. Finaly the gel was stained for 20 minutes in 1 x SYBR® Gold (Invitrogen) in TBE solution and washed with water. Stained nucleic acids were visulised on a Typhoon 9400 phosphoimager (Amersham).

3.4.9. Quantitative pcr analysis

Primer Mix

10 µl	forward primer 100 pM
10 µl	reverse primer 100 pM
320 µl	H_2O

PCR Master Mix (1x)

2.5 µl	Primer Mix
12.5 µl	Power Sybr® Green PCR Master Mix
5 µl	H_2O

Table 3.4 Oligonucleotides used for PCR

name	5'→3' sequence
Adh1-M_RT_For	AGCCGCTCACATTCCTCAAG
Adh1-M_RT_Rev	ACGGTGATACCAGCACACAAGA
Ref_RT_For	TGCGTACAAAAAGTGTCAAGAGATT
Ref_RT_Rev	ATGCGCAAGAAGGTGCCTAT

Table 3.5 PCR program details

Cycles No.	Name	Temp.	Duration
1		50	2 min
1		95	10 min
40		95	15 sec
		60	15 sec

For quantitative control of IP experiments as well as Whole Genome Amplification of DNA, a real-time PCR procedure was applied. All real-time PCR experiments were performed on ABI Prism 7000 Sequence Detection System using Power Sybr® Green PCR Master Mix (Applied Biosystems) with ROX as passive reference dye. All the samples were tested with two primer pairs listed in the Table 3.4.

In every the experiment, a standardization curve was made for each primer pair by using four serial dilutions (1:10) of a Sacharomyces Serevisiae genomic DNA stock solution (25 ng/µl). Tested probes included Input DNA (0.4 ng/µl), IP, mock IP and control IP with empty beads incubated with chromatin solution diluted 1:5 and 1:25. WGA procedure was monitored for introducing of unspecific bias by preferential amplification of some DNA pieces by testing Input and IP DNA after Amplification and Re-amplification steps. WGA amplified samples were diluted to 1 ng/µl upon usage. A control non-template (NTC) PCR reaction using water replacing DNA template was performed during each experiment to check for non-specific signal arising from primer dimmers, template or reagents contamination. 5 µl of template dilution were mixed with 20 µl of Master Mix and pipeted into MicroAmp™ Optical 96-well Reaction Plates (Applied Biosystems). Duplicates and triplicates of each genomic DNA dilution and each test probes, respectively, were performed. In order to avoid potential cross contamination of reagents DNase/RNase-free sterile water and filter barrier pipette

tips were used for all procedures. Table 3.5 shows the details of the PCR amplification program.

3.4.10. **Microarray handling**

Labeling, hybridization, array scanning, data extraction, and a preliminary data analysis were performed by imaGenes GmbH, as part of the NimbleGen ChIP-chip service (http://www.imagenes-bio.de/services/nimblegen/chip). For all experiments the NimbleGen *S. cerevisiae* ChIP Whole Genome Tiling Array (Cat.No. C4214-00-01) was used.   ChIP-chip datasets were derived from three or two biological replicates of S288C Rpb7-TAP or Rpb3-TAP strains, respectively, and from single experiments on W303 wild type and W303 Rpb3-TAP strains. All microarray measurements were performed with two-colour technology using Cy3 and Cy5, i.e. the DNA bound to immunoprecipitated protein was labelled with one dye and genomic background DNA with another dye, for a standard application the Input sample was labelled with Cy3 whereas IP sample with Cy5 dye. In case of two measurements, of S288CC Rpb3-TAP and S288CC Rpb7-TAP, respectively, the dye swapping, with IP sample labelled with Cy3 and Input sample labelled with Cy5 was performed.

3.5. **Bioinformatic analysis**

Bioinformatic analysis was performed by Holger Hartmann and Johannes Söding, Gene Center Munich, as a part of collaboration.

In all analyses, the logarithm of the fluorescent signal from the ChIP DNA divided by the signal from the genomic background was used. A standard background correction was performed on all such signals by subtracting their genome-wide average. The ChIP-chip measurements for TAP-tagged Rpb3 and Rpb7 were repeated with exchanged dyes Cy5 and Cy3 and averaged over measurements to subtract out the strong, systematic, dye-related technical noise. For Rpb3 two replica measurements with ChIP DNA labeled by Cy5, but only one with exchanged dyes was available. When averaging over the replicas, the Cy3-labeled signal was weighted doubly to cancel out systematic noise. To directly compare the Rpb3 and Rpb7 signals, their relative scales were adjusted, which can differ slightly from 1 due to technical effects (Do & Choi, 2006). For this purpose, a linear regression on the Rpb7 vs. Rpb3 scatter plot was performed.

The agreement between various pairs of genome-wide measurements of Rpb3 and Rpb7 occupancy was quantified by the Pearson correlation coefficient. Since the Pearson correlation coefficient is sensitive to the presence of noise, the coefficients

for the raw data and for smoothed traces were calculated. A local quadratic regression smoother with a Gaussian kernel ($\sigma=2$, Table 1) was used, corresponding to averaging over approximately nine data points, respectively. For analyzing correlations between occupancy signals and genomic features smoothed curves were used.

Abbreviations

ABC	subunits common for of Pol I, Pol II and Pol III
AC	subunits common for of Pol I and Pol III
C	subunit of Pol III (=RNA polymerase C)
C.elgans	*Caenorhabditis elegans*
ChIP	chromatin immunoprecipitacion
ChIP-chip	chromatin immunoprecipitacion coupled with a microarray
CTD	C-terminal domain of Rpb1 of Pol II
D.melanogaster	*Drosophila melanogaster*
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DSE	Distal sequence element
DTT	dithiothreitol
E.coli	*Escherichia coli*
EC	elongation complex
EDTA	ethylene diamine tetraacetic acid
EM	electron microscopy
EMSA	electrophoretic mobility shift assay
GTF	general transcription factor
Hepes	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRDC	helicase and RNAse D C-terminal domain (in Rpb4 and C17)
H.sapiens	*Homo sapiens*
ICR	internal control region
IP	immunoprecipitated
IPTG	isopropyl β-D-1-thiogalactopyranoside
miRNA	micro RNA
mRNA	messenger RNA
MWCO	molecular weight cutoff
NAC	nucleotide addition cycle
NTC	non template control
PCR	polymerase chain reaction
PDB	protein data bank
PEG	polyethylene glycol (number indicates average molecular weight in Da)
PIC	preinitiation complex

PMSF	phenylmethylsulfonyl fluoride
Pol	eukaryotic DNA-dependent RNA polymerase
PSE	proximal sequence element
RMSD	root mean square deviation
RNA	ribonucleic acid
Rpa	subunit of Pol I (=RNA polymerase A)
Rpb	subunit of Pol II (=RNA polymerase B)
Rpo	subunit of archeal RNA Polymerase
rRNA	ribosomal RNA
rt-PCR	real time pcr
sbRNA	stem-bulge RNA
S.cerevisiae	*Saccharomyces cerevisiae*
SINE	short interspersed repeated DNA elements
SNAPc	snRNA activating protein complex
snRNA	small nuclear RNA
snoRNA	Small nucleolar RNA
S.pombe	*Schizosaccharomyces pombe*
snoRNA	Small nucleolar RNA
TAF	TBP-associated factor
TAP tag	tandem affinity purification tag
TBP	TATA binding protein
TCEP	tris(2-carboxyethyl)phosphine
TFII	transcription factor of Pol II transcription
TFIII	transcription factor of Pol III transcription
TOR	Target of Rapamycin pathway
Tris	trishydroxymethylaminomethane
tRNA	transfer RNA
Tth	*Thermus thermophilus*
UTP	uridine triphosphate
vRNA	valut RNA
WGA	Whole genome amplification

References

Andrau J C, Sentenac A and Werner M. **1999.** Mutagenesis of yeast TFIIIB70 reveals C-terminal residues critical for interaction with TBP and C34. *J Mol Biol*, **288**:511-520.

Aparicio O, Geisberg J V, Sekinger E, Yang A, Moqtaderi Z and Struhl K. **2005.** Chromatin immunoprecipitation for determining the association of proteins with specific genomic sequences in vivo. *Curr Protoc Mol Biol*, **Chapter 21**:Unit 21 23.

Armache K-J, Kettenberger H and Cramer P. **2003.** Architecture of the initiation-competent 12-subunit RNA polymerase II. *Proc. Natl. Ac. Sc. USA*, **100**:6964-6968.

Armache K-J, Mitterweger S, Meinhart A and Cramer P. **2005.** Structures of complete RNA polymerase II and its subcomplex Rpb4/7. *J. Biol. Chem.*, **280**:7131-7134.

Armstrong J A. **2007.** Negotiating the nucleosome: factors that allow RNA polymerase II to elongate through chromatin. *Biochem Cell Biol*, **85**:426-434.

Asturias F, Meredith G, Poglitsch C and Kornberg R. **1997.** Two conformations of RNA polymerase II revealed by electron

crystallography. *J. Mol. Biol.*, **272**:536-540.

Baltimore D. **1970.** RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature*, **226**:1209-1211.

Bar-Nahum G, Epshtein V, Ruckenstein A E, Rafikov R, Mustaev A and Nudler E. **2005.** A ratchet mechanism of transcription elongation and its control. *Cell*, **120**:183-193.

Bartholomew B, Durkovich D, Kassavetis G A and Geiduschek E P. **1993.** Orientation and topography of RNA polymerase III in transcription complexes. *Mol Cell Biol*, **13**:942-952.

Bartholomew B, Kassavetis G A and Geiduschek E P. **1991.** Two components of Saccharomyces cerevisiae transcription factor IIIB (TFIIIB) are stereospecifically located upstream of a tRNA gene and interact with the second-largest subunit of TFIIIC. *Mol Cell Biol*, **11**:5181-5189.

Bhargava P and Kassavetis G A. **1999.** Abortive initiation by Saccharomyces cerevisiae RNA polymerase III. *J Biol Chem*, **274**:26550-26556.

Blencowe B J. **2002.** Transcription: surprising role for an elusive small nuclear RNA. *Curr Biol*, **12**:R147-149.

Bobkova E V and Hall B D. **1997.** Substrate specificity of the RNase activity of yeast RNA polymerase III. *J Biol Chem*, **272**:22832-22839.

Bogenhagen D F, Sakonju S and Brown D D. **1980.** A control region in the center of the 5S RNA gene directs specific initiation of transcription: II. The 3' border of the region. *Cell*, **19**:27-35.

Borchert G M, Lanier W and Davidson B L. **2006.** RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol*, **13**:1097-1101.

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

Bradfor M M. **1976.** A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, **72**:248-254.

Brow D A and Guthrie C. **1990.** Transcription of a yeast U6 snRNA gene requires a polymerase III promoter element in a novel position. *Genes Dev,* **4**:1345-1356.

Brueckner F, Armache K J, Cheung A, Damsma G E, Kettenberger H, Lehmann E, Sydow J F and Cramer P. **2008.** Structure-function studies of the RNA polymerase II elongation complex. *Acta Crystallographica Section D: Biological Crystallography*, **in press**.

Brueckner F and Cramer P. **2008.** Structural basis of transcription inhibition by alpha-amanitin and implications for RNA polymerase II translocation. *Nat Struct Mol Biol*, **15**:811-818.

Brun I, Sentenac A and Werner M. **1997.** Dual role of the C34 subunit of RNA polymerase III in transcription initiation. *Embo J*, **16**:5730-5741.

Budisa N, Steipe B, Demange P, Eckerskorn C, Kellermann J and Huber R. **1995.** High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in Escherichia coli. *Eur J Biochem*, **230**:788-796.

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

Bushnell D A, Cramer P and Kornberg R D. **2002.** Structural basis of transcription: alpha-amanitin-RNA polymerase II cocrystal at 2.8 A resolution. *Proc Natl Acad Sci U S A*, **99**:1218-1222.

Bushnell D A and Kornberg R D. **2003.** Complete RNA polymerase II at 4.1 Å resolution: implications for the initiation of transcription. *Proc Natl Acad Sci U S A*, **100**:6969-6972.

Bushnell D A, Westover K D, Davis R E and Kornberg R D. **2004.** Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms. *Science*, **303**:983-988.

Chedin S, Ferri M L, Peyroche G, Andrau J C, Jourdain S, Lefebvre O, Werner M, Carles C and Sentenac A. **1998.** The yeast RNA polymerase III transcription machinery: a paradigm for eukaryotic gene activation. *Cold Spring Harb Symp Quant Biol*, **63**:381-389.

Chedin S, Riva M, Schultz P, Sentenac A and Carles C. **1998.** The RNA cleavage activity of RNA polymerase III is mediated by an essential TFIIS-like subunit and is important for transcription termination. *Genes Dev*, **12**:3857-3871.

Chen B S, Mandal S S and Hampsey M. **2004.** High-resolution protein-DNA contacts for the yeast RNA polymerase II general transcription machinery. *Biochemistry*, **43**:12741-12749.

Chen H-T, Warfield L and Hahn S. **2007.** The positions of TFIIF and TFIIE in the RNA polymerase II transcription initiation complex. *Nat Struct Mol Biol*, **8**:696-703.

Choder M. **2004.** Rpb4 and Rpb7: subunits of RNA polymerase II and beyond. *Trends Biochem Sci*, **29**:674-681.

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

Ciesla M and Boguta M. **2008.** Regulation of RNA polymerase III transcription by Maf1 protein. *Acta Biochim Pol*, **55**:215-225.

Ciesla M, Towpik J, Graczyk D, Oficjalska-Pham D, Harismendy O, Suleau A, Balicki K, Conesa C, Lefebvre O and Boguta M. **2007.** Maf1 is involved in coupling carbon metabolism to RNA polymerase III transcription. *Mol Cell Biol*, **27**:7693-7702.

Cojocaru M, Jeronimo C, Forget D, Bouchard A, Bergeron D, Cote P, Poirier G G, Greenblatt J and Coulombe B. **2008.** Genomic location of the human RNA polymerase II general machinery: evidence for a role of TFIIF and Rpb7 at both early and late stages of transcription. *Biochem J*, **409**:139-147.

Conaway J W and Conaway R C. **1999.** Transcription elongation and human disease. *Annu Rev Biochem*, **68**:301-319.

Cozzarelli N R, Gerrard S P, Schlissel M, Brown D D and Bogenhagen D F. **1983.** Purified RNA polymerase III accurately and efficiently terminates transcription of 5S RNA genes. *Cell*, **34**:829-835.

Craighead J L, Chang W H and Asturias F J. **2002.** Structure of yeast RNA polymerase II in solution: implications for enzyme regulation and interaction with promoter DNA. *Struct. Fold. Des.*, **10**:1117-1125.

Cramer P. **2007.** Finding the right spot to start transcription. *Nat. Struct. Mol. Biol.*, **14**:686-687.

Cramer P, Armache K J, Baumli S, Benkert S, Brueckner F, Buchen C, Damsma G E, Dengl S, Geiger S R, Jasiak A J, Jawhari A, Jennebach S, Kamenski T, Kettenberger H, Kuhn C D, Lehmann E, Leike K, Sydow J F and Vannini A. **2008.** Structure of eukaryotic RNA polymerases. *Annu Rev Biophys*, **37**:337-352.

Cramer P, Bushnell D A, Fu J, Gnatt A L, Maier-Davis B, Thompson N E, Burgess R R, Edwards A M, David P R and Kornberg R D. **2000.** Architecture of RNA polymerase II and implications for the transcription mechanism [see comments]. *Science*, **288**:640-649.

Cramer P, Bushnell D A and Kornberg R D. **2001.** Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science*, **292**:1863-1876.

Crick F. **1970.** Central dogma of molecular biology. *Nature*, **227**:561-563.

Darst S A, Edwards A M, Kubalek E W and Kornberg R D. **1991.** Three-dimensional structure of yeast RNA polymerase II at 16Å resolution. *Cell*, **66**:121-128.

Das G, Henning D, Wright D and Reddy R. **1988.** Upstream regulatory elements are necessary and sufficient for transcription of a U6 RNA gene by RNA polymerase III. *Embo J*, **7**:503-512.

DeLano W L. **2002.** The PyMOL Molecular Graphics System. *DeLano Scientific*:San Carlos, CA, USA.

Deng W, Zhu X, Skogerbo G, Zhao Y, Fu Z, Wang Y, He H, Cai L, Sun H, Liu C, Li B, Bai B, Wang J, Jia D, Sun S, He H, Cui Y, Wang Y, Bu D and Chen R. **2006.** Organization of the Caenorhabditis elegans small non-coding transcriptome: genomic features, biogenesis, and expression. *Genome Res*, **16**:20-29.

Desai N, Lee J, Upadhya R, Chu Y, Moir R D and Willis I M. **2005.** Two steps in Maf1-dependent repression of transcription by RNA polymerase III. *J Biol Chem*, **280**:6455-6462.

Dieci G, Fiorino G, Castelnuovo M, Teichmann M and Pagano A. **2007.** The expanding RNA polymerase III transcriptome. *Trends Genet*, **23**:614-622.

Dieci G, Percudani R, Giuliodori S, Bottarelli L and Ottonello S. **2000.** TFIIIC-independent in vitro transcription of yeast tRNA genes. *J Mol Biol*, **299**:601-613.

Dieci G and Sentenac A. **1996.** Facilitated Recycling Pathway for RNA Polymerase III. *Cell*, **84**:245-252.

Dieci G and Sentenac A. **2003.** Detours and shortcuts to transcription reinitiation. *Trends in Biochemical Sciences*, **28**:202-209.

Do J H and Choi D K. **2006.** Normalization of microarray data: single-labeled and dual-labeled arrays. *Mol Cells*, **22**:254-261.

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

Engelke D R, Ng S Y, Shastry B S and Roeder R G. **1980.** Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. *Cell*, **19**:717-728.

Epshtein V, Mustaev A, Markovtsov V, Bereshchenko O, Nikiforov V and Goldfarb A. **2002.** Swing-gate model of nucleotide entry into the RNA polymerase active center. *Mol Cell*, **10**:623-634.

Eschenlauer J B, Kaiser M W, Gerlach V L and Brow D A. **1993.** Architecture of a yeast U6 RNA gene promoter. *Mol Cell Biol*, **13**:3015-3026.

Fernandez-Tornero C, Bottcher B, Riva M, Carles C, Steuerwald U, Ruigrok R W, Sentenac A, Muller C W and Schoehn G. **2007.** Insights into Transcription Initiation and Termination from the Electron Microscopy Structure of Yeast RNA Polymerase III. *Mol Cell*, **25**:813-823.

Ferrari R, Rivetti C, Acker J and Dieci G. **2004.** Distinct roles of transcription factors TFIIIB and TFIIIC in RNA polymerase III transcription reinitiation. *Proc Natl Acad Sci U S A*, **101**:13442-13447.

Ferri M L, Peyroche G, Siaut M, Lefebvre O, Carles C, Conesa C and Sentenac A. **2000.** A novel subunit of yeast RNA polymerase III interacts with the TFIIB-related domain of TFIIIB70. *Mol Cell Biol*, **20**:488-495.

Fontana A, de Laureto P P, Spolaore B, Frare E, Picotti P and Zambonin M. **2004.** Probing protein structure by limited proteolysis. *Acta Biochim Pol*, **51**:299-321.

Fontana A, Fassina G, Vita C, Dalzoppo D, Zamai M and Zambonin M. **1986.** Correlation between sites of limited proteolysis and segmental mobility in thermolysin. *Biochemistry*, **25**:1847-1851.

French S L, Osheim Y N, Schneider D A, Sikes M L, Fernandez C F, Copela L A, Misra V A, Nomura M, Wolin S L and Beyer A L. **2008.** Visual analysis of the yeast 5S rRNA gene transcriptome: regulation and role of La protein. *Mol Cell Biol*, **28**:4576-4587.

Fu J, Gnatt A L, Bushnell D A, Jensen G, J,, Thompson N E, Burgess R R, David P R and Kornberg R D. **1999.** Yeast RNA polymerase II at 5 Å resolution. *Cell*, **98**:799-810.

Geiduschek E P and Kassavetis G A. **2001.** The RNA polymerase III transcription apparatus. *J Mol Biol*, **310**:1-26.

Geiger S R, Kuhn C D, Leidig C, Renkawitz J and Cramer P. **2008.** Crystallization of RNA polymerase I subcomplex A14/A43 by iterative prediction, probing and removal of flexible regions. *Acta Crystallogr Sect F Struct Biol Cryst Commun*, **64**:413-418.

Gilmour D S and Fan R. **2008.** Derailing the locomotive: transcription termination. *J Biol Chem*, **283**:661-664.

Giuliodori S, Percudani R, Braglia P, Ferrari R, Guffanti E, Ottonello S and Dieci G. **2003.** A composite upstream sequence motif potentiates tRNA gene transcription in yeast. *J Mol Biol*, **333**:1-20.

Gnatt A L, Cramer P, Fu J, Bushnell D A and Kornberg R D. **2001.** Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. *Science*, **292**:1876-1882.

Gomez-Roman N, Grandori C, Eisenman R N and White R J. **2003.** Direct activation of RNA polymerase III transcription by c-Myc. *Nature*, **421**:290-294.

Gong X Q, Nedialkov Y A and Burton Z F. **2004.** Alpha-amanitin blocks translocation by human RNA polymerase II. *J Biol Chem*, **279**:27422-27427.

Gupta S and Reddy R. **1991.** Compilation of small RNA sequences. *Nucleic Acids Res*, **19 Suppl**:2073-2075.

Gusarov I and Nudler E. **1999.** The mechanism of intrinsic transcription termination. *Mol Cell*, **3**:495-504.

Hahn S. **2004.** Structure and mechanism of the RNA polymerase II transcription machinery. *Nat Struct Mol Biol*, **11**:394-403.

Hallberg E M, Fung P and Hallberg R L. **1992.** Genomic sequence encoding a heat shock-induced, RNA polymerase III-transcribed RNA from Tetrahymena thermophila. *Nucleic Acids Res*, **20**:912.

Hanahan D and Weinberg R A. **2000.** The hallmarks of cancer. *Cell*, **100**:57-70.

Herr A J, Jensen M B, Dalmay T and Baulcombe D C. **2005.** RNA polymerase IV directs silencing of endogenous DNA. *Science*, **308**:118-120.

Hirata A, Klein B J and Murakami K S. **2008.** The X-ray crystal structure of RNA polymerase from Archaea. *Nature*, **451**:851-854.

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

Hsieh Y J, Kundu T K, Wang Z, Kovelman R and Roeder R G. **1999.** The TFIIIC90 subunit of TFIIIC interacts with multiple components of the RNA polymerase III machinery and contains a histone-specific acetyltransferase activity. *Mol Cell Biol*, **19**:7697-7704.

Hu P, Wu S, Sun Y, Yuan C C, Kobayashi R, Myers M P and Hernandez N. **2002.** Characterization of human RNA polymerase III identifies orthologues for Saccharomyces cerevisiae RNA polymerase III subunits. *Mol Cell Biol*, **22**:8044-8055.

Huang Y and Maraia R J. **2001.** Comparison of the RNA polymerase III transcription machinery in Schizosaccharomyces pombe, Saccharomyces cerevisiae and human. *Nucleic Acids Res*, **29**:2675-2690.

Isogai Y, Takada S, Tjian R and Keles S. **2007.** Novel TRF1/BRF target genes revealed by genome-wide analysis of Drosophila Pol III transcription. *Embo J*, **26**:79-89.

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

Jasiak A J, Armache K J, Martens B, Jansen R P and Cramer P. **2006.** Structural biology of RNA polymerase III: subcomplex C17/25 X-ray structure and 11 subunit enzyme model. *Mol Cell*, **23**:71-81.

Jensen G J, Meredith G, Bushnell D A and Kornberg R D. **1998.** Structure of wild-type yeast RNA polymerase II and location of Rpb4 and Rpb7. *EMBO J.*, **17**:2353-2358.

Jones T A, Zou J Y, Cowan S W and Kjeldgaard M. **1991.** Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst.*, **A47**:110-119.

Kanno T, Huettel B, Mette M F, Aufsatz W, Jaligot E, Daxinger L, Kreil D P, Matzke M and Matzke A J. **2005.** Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nat Genet*, **37**:761-765.

Kaplan C D, Larsson K M and Kornberg R D. **2008.** The RNA polymerase II trigger loop functions in substrate selection and is directly targeted by alpha-amanitin. *Mol Cell*, **30**:547-556.

Kassavetis G A, Bartholomew B, Blanco J A, Johnson T E and Geiduschek E P. **1991.** Two essential components of the Saccharomyces cerevisiae transcription factor TFIIIB: transcription and DNA-binding properties. *Proc Natl Acad Sci U S A*, **88**:7308-7312.

Kassavetis G A, Blanco J A, Johnson T E and Geiduschek E P. **1992.** Formation of open and elongating transcription complexes by RNA polymerase III. *Journal of Molecular Biology*, **226**:47-58.

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.

Kettenberger H, Armache K-J and Cramer P. **2004.** Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. *Mol. Cell*, **16**:955-965.

Khazak V, Estojak J, Cho H, Majors J, Sonoda G, Testa J R and Golemis E A. **1998.** Analysis of the interaction of the novel RNA polymerase II (pol II) subunit hsRPB4 with its partner hsRPB7 and with pol II. *Mol Cell Biol*, **18**:1935-1945.

Khoo B, Brophy B and Jackson S P. **1994.** Conserved functional domains of the RNA polymerase III general transcription factor BRF. *Genes Dev*, **8**:2879-2890.

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

Kuhn C D, Geiger S R, Baumli S, Gartmann M, Gerber J, Jennebach S, Mielke T, Tschochner H, Beckmann R and Cramer P. **2007.** Functional architecture of RNA polymerase I. *Cell*, **131**:1260-1272.

Kulaeva O I, Gaykalova D A and Studitsky V M. **2007.** Transcription through chromatin by RNA polymerase II: histone displacement and exchange. *Mutat Res*, **618**:116-129.

Kunkel G R and Pederson T. **1988.** Upstream elements required for efficient transcription of a human U6 RNA gene resemble those of U1 and U2 genes even though a different polymerase is used. *Genes Dev*, **2**:196-204.

Kusser A G, Bertero M G, Naji S, Becker T, Thomm M, Beckmann R and Cramer P. **2008.** Structure of an archaeal RNA polymerase. *J Mol Biol*, **376**:303-307.

Laemmli U K. **1970.** Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**:680-685.

Landrieux E, Alic N, Ducrot C, Acker J, Riva M and Carles C. **2006.** A subcomplex of RNA polymerase III subunits involved in transcription termination and reinitiation. *Embo J*, **25**:118-128.

Lassar A B, Martin P L and Roeder R G. **1983.** Transcription of class III genes: formation of preinitiation complexes. *Science*, **222**:740-748.

Li B, Carey M and Workman J L. **2007.** The role of chromatin during transcription. *Cell*, **128**:707-719.

Liu C L, Kaplan T, Kim M, Buratowski S, Schreiber S L, Friedman N and Rando O J. **2005.** Single-nucleosome mapping of histone modifications in S-cerevisiae. *Plos Biology*, **3**:1753-1769.

Lorenzen K, Vannini A, Cramer P and Heck A J. **2007.** Structural biology of RNA polymerase III: mass spectrometry elucidates subcomplex architecture. *Structure*, **15**:1237-1245.

Lotan R, Bar-On V G, Harel-Sharvit L, Duek L, Melamed D and Choder M. 2005. The RNA polymerase II subunit Rpb4p mediates decay of a specific class of mRNAs. *Genes Dev*, **19**:3004-3016.

Maillet I, Buhler J M, Sentenac A and Labarre J. 1999. Rpb4p is necessary for RNA polymerase II activity at high temperature. *J Biol Chem*, **274**:22586-22590.

Martignetti J A and Brosius J. 1993. BC200 RNA: a neural RNA polymerase III product encoded by a monomeric Alu element. *Proc Natl Acad Sci U S A*, **90**:11563-11567.

Martignetti J A and Brosius J. 1995. BC1 RNA: transcriptional analysis of a neural cell-specific RNA polymerase III transcript. *Mol Cell Biol*, **15**:1642-1650.

Matsuzaki H, Kassavetis G A and Geiduschek E P. 1994. Analysis of RNA Chain Elongation and Termination by Saccharomyces cerevisiae RNA Polymerase III. *Journal of Molecular Biology*, **235**:1173-1192.

Mattick J S and Makunin I V. 2006. Non-coding RNA. *Hum Mol Genet*, **15 Spec No 1**:R17-29.

McCoy A J, Grosse-Kunstleve R W, Storoni L C and Read R J. 2005. Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr*, **61**:458-464.

Meinhart A, Blobel J and Cramer P. 2003. An extended winged helix domain in general transcription factor E/IIEalpha. *J Biol Chem*, **278**:48267-48274.

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

Meka H, Daoust G, Arnvig K B, Werner F, Brick P and Onesti S. 2003. Structural and functional homology between the RNAP(I) subunits A14/A43 and the archaeal RNAP subunits E/F. *Nucleic Acids Res*, **31**:4391-4400.

Meka H, Werner F, Cordell S C, Onesti S and Brick P. 2005. Crystal structure and RNA binding of the Rpb4/Rpb7 subunits of human RNA polymerase II. *Nucleic Acids Res*, **33**:6435-6444.

Miller G and Hahn S. 2006. A DNA-tethered cleavage probe reveals the path for promoter DNA in the yeast preinitiation complex. *Nat Struct Mol Biol*, **13**:603-610.

Minakhin L, Bhagat S, Brunning A, Campbell E A, Darst S A, Ebright R H and Severinov K. 2001. Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. *Proc Natl Acad Sci U S A*, **98**:892-897.

Mitchell M T, Hobson G M and Benfield P A. 1992. TATA box-mediated polymerase III transcription in vitro. *J Biol Chem*, **267**:1995-2005.

Mitsuzawa H, Kanda E and Ishihama A. 2003. Rpb7 subunit of RNA polymerase II interacts with an RNA-binding protein involved in processing of transcripts. *Nucleic Acids Res*, **31**:4696-4701.

Miyao T, Barnett J D and Woychik N A. 2001. Deletion of the RNA polymerase subunit RPB4 acts as a global, not stress-specific, shut-off switch for RNA

polymerase II transcription at high temperatures. *J Biol Chem*, **276**:46408-46413.

Moir R D, Lee J, Haeusler R A, Desai N, Engelke D R and Willis I M. **2006.** Protein kinase A regulates RNA polymerase III transcription through the nuclear localization of Maf1. *Proc Natl Acad Sci U S A*, **103**:15044-15049.

Morozov V, Mushegian A R, Koonin E V and Bork P. **1997.** A putative nucleic acid-binding domain in Bloom's and Werner's syndrome helicases. *Trends Biochem Sci*, **22**:417-418.

Mrazek J, Kreutmayer S B, Grasser F A, Polacek N and Huttenhofer A. **2007.** Subtractive hybridization identifies novel differentially expressed ncRNA species in EBV-infected human B cells. *Nucl. Acids Res.*, **35**:e73-.

Murphy S, Tripodi M and Melli M. **1986.** A sequence upstream from the coding region is required for the transcription of the 7SK RNA genes. *Nucleic Acids Res*, **14**:9243-9260.

Mus E, Hof P R and Tiedge H. **2007.** Dendritic BC200 RNA in aging and in Alzheimer's disease. *Proc Natl Acad Sci U S A*, **104**:10679-10684.

Nakaya H I, Amaral P P, Louro R, Lopes A, Fachel A A, Moreira Y B, El-Jundi T A, da Silva A M, Reis E M and Verjovski-Almeida S. **2007.** Genome mapping and expression analyses of human intronic noncoding RNAs reveal tissue-specific patterns and enrichment in genes related to regulation of transcription. *Genome Biol*, **8**:R43.

Oficjalska-Pham D, Harismendy O, Smagowicz W J, Gonzalez de Peredo A, Boguta M, Sentenac A and Lefebvre O. **2006.** General Repression of RNA Polymerase III Transcription Is Triggered by Protein Phosphatase Type 2A-Mediated Dephosphorylation of Maf1. *Molecular Cell*, **22**:623-632.

Orlicky S M, Tran P T, Sayre M H and Edwards A M. **2001.** Dissociable Rpb4-Rpb7 subassembly of rna polymerase II binds to single- strand nucleic acid and mediates a post-recruitment step in transcription initiation. *J Biol Chem*, **276**:10097-10102.

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.

Otwinowski Z and Minor W. **1996.** Processing of X-ray diffraction data collected in oscillation mode. *Meth. Enzym.*, **276**:307-326.

Ouyang C, Martinez M J, Young L S and Sprague K U. **2000.** TATA-Binding protein-TATA interaction is a key determinant of differential transcription of silkworm constitutive and silk gland-specific tRNA(Ala) genes. *Mol Cell Biol*, **20**:1329-1343.

Pagano A, Castelnuovo M, Tortelli F, Ferrari R, Dieci G and Cancedda R. **2007.** New small nuclear RNA gene-like transcriptional units as sources of regulatory transcripts. *PLoS Genet*, **3**:e1.

Parrott A M and Mathews M B. **2007.** Novel rapidly evolving hominid RNAs bind nuclear factor 90 and display tissue-restricted distribution. *Nucleic Acids Res*, **35**:6249-6258.

Peyroche G, Levillain E, Siaut M, Callebaut I, Schultz P, Sentenac A, Riva M and Carles C. **2002.** The A14-A43 heterodimer subunit in yeast RNA pol I and their relationship to Rpb4-Rpb7 pol II subunits. *Proc Natl Acad Sci U S A*, **99**:14670-14675.

Pfeffer S, Sewer A, Lagos-Quintana M, Sheridan R, Sander C, Grasser F A, van Dyk L F, Ho C K, Shuman S, Chien M, Russo J J, Ju J, Randall G, Lindenbach B D, Rice C M, Simon V, Ho D D, Zavolan M and Tuschl T. **2005.** Identification of microRNAs of the herpesvirus family. *Nat Methods*, **2**:269-276.

Pillai B, Sampath V, Sharma N and Sadhale P. **2001.** Rpb4, a non-essential subunit of core RNA polymerase II of Saccharomyces cerevisiae is important for activated transcription of a subset of genes. *J Biol Chem*, **276**:30641-30647.

Pillai B, Verma J, Abraham A, Francis P, Kumar Y, Tatu U, Brahmachari S K and Sadhale P P. **2003.** Whole genome expression profiles of yeast RNA polymerase II core subunit, Rpb4, in stress and nonstress conditions. *J Biol Chem*, **278**:3339-3346.

Pluta K, Lefebvre O, Martin N C, Smagowicz W J, Stanford D R, Ellis S R, Hopper A K, Sentenac A and Boguta M. **2001.** Maf1p, a negative effector of RNA polymerase III in Saccharomyces cerevisiae. *Mol Cell Biol*, **21**:5031-5040.

Proshkina G M, Shematorova E K, Proshkin S A, Zaros C, Thuriaux P and Shpakovski G V. **2006.** Ancient origin, functional conservation and fast evolution of DNA-dependent RNA polymerase III. *Nucleic Acids Res*, **34**:3615-3624.

Radonjic M, Andrau J C, Lijnzaad P, Kemmeren P, Kockelkorn T T, van Leenen D, van Berkum N L and Holstege F C. **2005.** Genome-wide analyses reveal RNA polymerase II located upstream of genes poised for rapid response upon S. cerevisiae stationary phase exit. *Mol Cell*, **18**:171-183.

Reina J H, Azzouz T N and Hernandez N. **2006.** Maf1, a new player in the regulation of human RNA polymerase III transcription. *PLoS ONE*, **1**:e134.

Reines D. **1992.** Elongation factor-dependent transcript shortening by template-engaged RNA polymerase II. *J Biol Chem*, **267**:3795-3800.

Ridanpaa M, van Eenennaam H, Pelin K, Chadwick R, Johnson C, Yuan B, vanVenrooij W, Pruijn G, Salmela R, Rockas S, Makitie O, Kaitila I and de la Chapelle A. **2001.** Mutations in the RNA component of RNase MRP cause a pleiotropic human disease, cartilage-hair hypoplasia. *Cell*, **104**:195-203.

Ridanpaa M, Ward L M, Rockas S, Sarkioja M, Makela H, Susic M, Glorieux F H, Cole W G and Makitie O. **2003.** Genetic changes in the RNA components of RNase MRP and RNase P in Schmid metaphyseal chondrodysplasia. *J Med Genet*, **40**:741-746.

Roberts D N, Wilson B, Huff J T, Stewart A J and Cairns B R. **2006.** Dephosphorylation and Genome-Wide Association of Maf1 with Pol III-Transcribed Genes during Repression. *Molecular Cell*, **22**:633-644.

Roeder R. **1996.** The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem. Sci.*, **21**:327-335.

Roeder R G. **1998.** Role of general and gene-specific cofactors in the regulation of eukaryotic transcription. *Cold Spring Harb Symp Quant Biol*, **63**:201-218.

Roifman C M, Gu Y and Cohen A. **2006.** Mutations in the RNA component of RNase mitochondrial RNA processing might cause Omenn syndrome. *J Allergy Clin Immunol*, **117**:897-903.

Rollins J, Veras I, Cabarcas S, Willis I and Schramm L. **2007.** Human Maf1 negatively regulates RNA polymerase III transcription via the TFIIB family members Brf1 and Brf2. *Int J Biol Sci*, **3**:292-302.

Romero D P and Blackburn E H. **1981.** A conserved secondary structure for telomerase RNA. *Cell*, **67**:343-353.

Rosenheck S and Choder M. **1998.** Rpb4, a subunit of RNA polymerase II, enables the enzyme to transcribe at temperature extremes in vitro. *J Bacteriol*, **180**:6187-6192.

Ruet A, Sentenac A, Fromageot P, Winsor B and Lacroute F. **1980.** A mutation of the B220 subunit gene affects the structural and functional properties of yeast RNA polymerase B in vitro. *J Biol Chem*, **255**:6450-6455.

Runner V M, Podolny V and Buratowski S. **2008.** The Rpb4 subunit of RNA polymerase II contributes to cotranscriptional recruitment of 3' processing factors. *Mol Cell Biol*, **28**:1883-1891.

Sadhale P P and Woychik N A. **1994.** C25, an essential RNA polymerase III subunit related to the RNA polymerase II subunit RPB7. *Mol Cell Biol*, **14**:6164-6170.

Sakonju S, Bogenhagen D F and Brown D D. **1980.** A control region in the center of the 5S RNA gene directs specific initiation of transcription: I. The 5' border of the region. *Cell*, **19**:13-25.

Sakonju S, Brown D D, Engelke D, Ng S Y, Shastry B S and Roeder R G. **1981.** The binding of a transcription factor to deletion mutants of a 5S ribosomal RNA gene. *Cell*, **23**:665-669.

Sakurai H and Ishihama A. **1997.** Gene organization and protein sequence of the small subunits of Schizosaccharomyces pombe RNA polymerase II. *Gene*, **196**:165-174.

Sakurai H, Mitsuzawa H, Kimura M and Ishihama A. **1999.** The Rpb4 subunit of fission yeast Schizosaccharomyces pombe RNA polymerase II is essential for cell viability and similar in structure to the corresponding subunits of higher eukaryotes. *Mol Cell Biol*, **19**:7511-7518.

Sambrook J and Russell D W. **2001.** *Molecular cloning – A laboratory Manual*, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Schramm L and Hernandez N. **2002.** Recruitment of RNA polymerase III to its target promoters. *Genes Dev*, **16**:2593-2620.

Schramm L, Pendergrast P S, Sun Y and Hernandez N. **2000.** Different human TFIIIB activities direct RNA polymerase III transcription from TATA-containing and TATA-less promoters. *Genes Dev*, **14**:2650-2663.

Scott P H, Cairns C A, Sutcliffe J E, Alzuherri H M, McLees A, Winter A G and White R J. **2001.** Regulation of RNA polymerase III transcription during cell cycle entry. *J Biol Chem*, **276**:1005-1014.

Selitrennik M, Duek L, Lotan R and Choder M. **2006.** Nucleocytoplasmic shuttling of the Rpb4p and Rpb7p subunits of Saccharomyces cerevisiae RNA polymerase II by two pathways. *Eukaryot Cell*, **5**:2092-2103.

Shpakovski G V and Shematorova E K. **1999.** Characterization of the rpa43+ cDNA of Schizosaccharomyces pombe: Structural similarity of subunit Rpa43 of RNA polymerase I and subunit Rpc25 of RNA polymerase III. *Russ. J. Bioorgan. Chem.*, **25**:700-704.

Siaut M, Zaros C, Levivier E, Ferri M L, Court M, Werner M, Callebaut I, Thuriaux P, Sentenac A and Conesa C. **2003.** An Rpb4/Rpb7-Like Complex in Yeast RNA Polymerase III Contains the Orthologue of Mammalian CGRP-RCP. *Mol Cell Biol*, **23**:195-205.

Specht T, Wolters J and Erdmann V A. **1991.** Compilation of 5S rRNA and 5S rRNA gene sequences. *Nucleic Acids Res*, **19 Suppl**:2189-2191.

Sprinzl M, Dank N, Nock S and Schon A. **1991.** Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res*, **19 Suppl**:2127-2171.

Steinberg T H, Mathews D E, Durbin R D and Burgess R R. **1990.** Tagetitoxin: a new inhibitor of eukaryotic transcription by RNA polymerase III. *J Biol Chem*, **265**:499-505.

Steinmetz E J, Warren C L, Kuehner J N, Panbehi B, Ansari A Z and Brow D A. **2006.** Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. *Mol Cell*, **24**:735-746.

Steitz T A. **1998.** A mechanism for all polymerases [news; comment]. *Nature*, **391**:231-232.

Studier F W. **2005.** Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif*, **41**:207-234.

Sussman D J, Chung J and Leder P. **1991.** In vitro and in vivo analysis of the c-myc RNA polymerase III promoter. *Nucleic Acids Res*, **19**:5045-5052.

Svejstrup J Q. **2004.** The RNA polymerase II transcription cycle: cycling through chromatin. *Biochim Biophys Acta*, **1677**:64-73.

Tan Q, Li X, Sadhale P P, Miyao T and Woychik N A. **2000.** Multiple mechanisms of suppression circumvent transcription defects in an RNA polymerase mutant. *Mol Cell Biol*, **20**:8124-8133.

Teichmann M, Wang Z and Roeder R G. **2000.** A stable complex of a novel transcription factor IIB- related factor, human TFIIB50, and associated proteins mediate selective transcription by RNA polymerase III of genes with upstream promoter elements. *Proc Natl Acad Sci U S A*, **97**:14200-14205.

Temin H M and Mizutani S. **1970.** RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature*, **226**:1211-1213.

Terwilliger T C. **2002.** Automated structure solution, density modification and model building. *Acta Crystallogr D Biol Crystallogr*, **58**:1937-1940.

Thieffry D and Sarkar S. **1998.** Forty years under the central dogma. *Trends Biochem Sci*, **23**:312-316.

Thompson J D, Higgins D G and Gibson T J. **1994.** CLUSTAL W: improving the sensibility of progressive multiple sequence alignment through sequence weighing, positions-specific gap penalties and weight matrix choice. *Nuc. Acid. Res.*, **22**:4673-4680.

Thuillier V, Stettler S, Sentenac A, Thuriaux P and Werner M. **1995.** A mutation in the C31 subunit of Saccharomyces cerevisiae RNA polymerase III affects transcription initiation. *Embo J*, **14**:351-359.

Till S and Ladurner A G. **2007.** RNA Pol IV plays catch with Argonaute 4. *Cell*, **131**:643-645.

Todone F, Brick P, Werner F, Weinzierl R O and Onesti S. **2001.** Structure of an archaeal homolog of the eukaryotic RNA polymerase II RPB4/RPB7 complex. *Mol Cell*, **8**:1137-1143.

Towpik J, Graczyk D, Gajda A, Lefebvre O and Boguta M. **2008.** Derepression of RNA polymerase III transcription by phosphorylation and nuclear export of its negative regulator, Maf1. *J Biol Chem*, **283**:17168-17174.

Ujvari A and Luse D S. **2006.** RNA emerging from the active site of RNA polymerase II interacts with the Rpb7 subunit. *Nat Struct Mol Biol*, **13**:49-54.

Upadhya R, Lee J and Willis I M. **2002.** Maf1 Is an Essential Mediator of Diverse Signals that Repress RNA Polymerase III Transcription. *Molecular Cell*, **10**:1489-1494.

van Zon A, Mossink M H, Scheper R J, Sonneveld P and Wiemer E A. **2003.** The vault complex. *Cell Mol Life Sci*, **60**:1828-1837.

Vassylyev D G, Sekine S, Laptenko O, Lee J, Vassylyeva M N, Borukhov S and Yokoyama S. **2002.** Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution. *Nature*, **417**:712-719.

Vassylyev D G, Vassylyeva M N, Perederina A, Tahirov T H and Artsimovitch I. **2007.** Structural basis for transcription elongation by bacterial RNA polymerase. *Nature*, **448**:157-162.

Vassylyev D G, Vassylyeva M N, Zhang J, Palangat M, Artsimovitch I and Landick R. **2007.** Structural basis for substrate loading in bacterial RNA polymerase. *Nature*, **448**:163-168.

Verma-Gaur J, Rao S N, Taya T and Sadhale P. **2008.** Genomewide Recruitment Analysis of Rpb4, a Subunit of Polymerase II in Saccharomyces cerevisiae, Reveals Its Involvement in Transcription Elongation. *Eukaryotic Cell*, **7**:1009-1018.

Wade J T and Struhl K. **2008.** The transition from transcriptional initiation to elongation. *Curr Opin Genet Dev*, **18**:130-136.

Wang D, Bushnell D A, Westover K D, Kaplan C D and Kornberg R D. **2006.** Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell*, **127**:941-954.

Wang D and Hawley D K. **1993.** Identification of a 3' - 5' exonucldase activity associate with human RNA polymerase II. *Proc. Natl. Acad. Sci. USA*, **90**:843-847.

Wang Z and Roeder R G. **1997.** Three human RNA polymerase III-specific subunits form a subcomplex with a selective function in specific transcription initiation. *Genes Dev*, **11**:1315-1326.

Warner J R. **1999.** The economics of ribosome biosynthesis in yeast. *Trends Biochem Sci*, **24**:437-440.

Weeks C M and Miller R. **1999.** Optimizing Shake-and-Bake for proteins. *Acta Crystallogr D Biol Crystallogr*, **55 (Pt 2)**:492-500.

Weinmann R and Roeder R G. **1974.** Role of DNA-dependent RNA polymerase 3 in the transcription of the tRNA and 5S RNA genes. *Proc Natl Acad Sci U S A*, **71**:1790-1794.

Werner M, Chaussivert N, Willis I M and Sentenac A. **1993.** Interaction between a complex of RNA polymerase III subunits and the 70-kDa component of transcription factor IIIB. *J Biol Chem*, **268**:20721-20724.

Werner M, Hermann-Le Denmat S, Treich I, Sentenac A and Thuriaux P. **1992.** Effect of mutations in a zinc-binding domain of yeast RNA polymerase C (III) on enzyme function and subunit association. *Mol Cell Biol*, **12**:1087-1095.

Westover K D, Bushnell D A and Kornberg R D. **2004.** Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center. *Cell*, **119**:481-489.

White R J. **1998.** *RNA Pol III Transcription*: Springer-Verlag.

White R J. **2004.** RNA polymerase III transcription and cancer. *Oncogene*, **23**:3208-3216.

White R J. **2005.** RNA polymerases I and III, growth control and cancer. *Nat Rev Mol Cell Biol*, **6**:69-78.

Willis I M. **1993.** RNA polymerase III. Genes, factors and transcriptional specificity. *Eur J Biochem*, **212**:1-11.

Willis I M and Moir R D. **2007.** Integration of nutritional and stress signaling pathways by Maf1. *Trends in Biochemical Sciences*, **32**:51-53.

Winter A G, Sourvinos G, Allison S J, Tosh K, Scott P H, Spandidos D A and White R J. **2000.** RNA polymerase III transcription factor TFIIIC2 is overexpressed in ovarian tumors. *Proc Natl Acad Sci U S A*, **97**:12619-12624.

Wolin S L. **1985.** Small cytoplasmic ribonucleoproteins. *Trends in Genetics*, **1**:201-204.

Woychik N A and Young R A. **1989.** RNA polymerase II subunit RPB4 is essential for high- and low-termperature yeast cell growth. *Mol. Cell. Biol.*, **9**:2854-2859.

Yang Y H and Speed T. **2002.** Design issues for cDNA microarray experiments. *Nat Rev Genet*, **3**:579-588.

Yukawa Y, Sugita M, Choisne N, Small I and Sugiura M. **2000.** The TATA motif, the CAA motif and the poly(T) transcription termination motif are all important for transcription re-initiation on plant tRNA genes. *Plant J*, **22**:439-447.

Zalfa F, Giorgi M, Primerano B, Moro A, Di Penta A, Reis S, Oostra B and Bagni C. **2003.** The fragile X syndrome protein FMRP associates with BC1 RNA and regulates the translation of specific mRNAs at synapses. *Cell*, **112**:317-327.

Zaros C and Thuriaux P. **2005.** Rpc25, a conserved RNA polymerase III subunit, is critical for transcription initiation. *Mol Microbiol*, **55**:104-114.

Zhang G, Campbell E A, Minakhin L, Richter C, Severinov K and Darst S A. **1999.** Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 A resolution [see comments]. *Cell*, **98**:811-824.

Curriculum vitae – Anna Justyna Jasiak

1. Personal Details

Name	Anna Justyna Jasiak
Institute address	Department of Chemistry and Biochemistry
	Gene Center
	University of Munich
	Feodor-Lynen-Str. 25
	D-81377 München
Institute phone	+49-89-2180-76955
Email	jasiak@lmb.uni-muenchen.de
Date of birth	November 22, 1980
Place of birth	Kedzierzyn-Kozle, Poland
Nationality	Polish
Languages	Polish – native speaker
	English – excellent command
	German - very good command

2. Curriculum vitae

1987-1995   **Grammar school** in Kedzierzyn-Kozle, Poland

06/1999   *Matura* (Polish equivalent to A-levels)

1999-2004   **MSc degree in Molecular Biology**

                       Jagiellonian University of Cracow, Poland

                       Specialization in physical biochemistry.

07/2003   **Intership** in Laboratory of Molecular Biology of
                       Microorganisms, Institute of Immunology and
                       Experimental Therapy, Polish Academy of Sciences,
                       Wroclaw, Poland

10/2003 – 04/2004   **Intership** in Department of Structure Research, Max-
                       Planck-Institute of Biochemistry, Martinsried, Germany

10/2004 - present   **Ph.D studies**, Prof. Patrick Cramer Group,
                       Gene Center, LMU Munich, Germany

                       Gene transcription and gene regulation

                       *Topic: Structure-function analysis of the RNA
                       polymerase III subcomplex C17/25 and genome-wide
                       distribution of RNA polymerase II*

2004 - 2005   **Doctoral scholarship** of Elite Network of Bavaria

2004 - present   **Member** of the International Graduate School

                       „Nano-Bio-Technology", Center of NanoScience,

                       LMU Munich

01/2009   **PhD thesis defense**

3. Conferences

2005	„NanoScience: From Quantum Bits to Life Science" Center of NanoScience Winterschool
	Mauterndorf, Austria
	Participant
	56. Mosbacher Kolloquium der Gesellschaft für Biochemie und Molekularbiologie „Molecular Machines"
	Mosbach, Germany
	Participant
2006	ESF Research Conference „Gene Transcription in Yeast"
	Sant Feliu de Guixols, Spain
	Poster presentation
2007	32nd FEBS Congerss „Molecular Machines"
	Vienna, Austria
	Poster presentation
	2nd Transregio 5 Symposium Chromatin „Assembly and Inheritance of Functional States"
	Munich, Germany
	Participant