

---

Dissertation zur Erlangung des Doktorgrades der Fakultät für Chemie und Pharmazie  
der Ludwig-Maximilians-Universität München

**The Mediator head module and  
regulation of RNA polymerase II  
transcription initiation**

Martin Josef Seizl  
aus München  
2011

---

## **Erklärung**

Diese Dissertation wurde im Sinne von §13 Abs. 3 der Promotionsordnung vom 29. Januar 1998 (in der Fassung der vierten Änderungssatzung vom 26. November 2004) von Herrn Prof. Dr. Patrick Cramer betreut.

## **Ehrenwörtliche Versicherung**

Diese Dissertation wurde selbstständig und ohne unerlaubte Hilfe erarbeitet.

München,

---

Martin Josef Seizl

Dissertation eingereicht am	07.03.2011
1. Gutachter	Prof. Dr. Patrick Cramer
2. Gutachter	Prof. Dr. Dietmar Martin
Mündliche Prüfung am	04.04.2011

## Acknowledgements

*“Coming together is a beginning. Keeping together is progress.*

*Working together is success.”*

(Henry Ford, 1863-1947)

This quote describes very well the multidisciplinary and collaborative atmosphere at the Gene Center in Munich. Therefore I am very glad to have decided to do my PhD thesis in the laboratory of Prof. Patrick Cramer.

First of all, I would like to thank my supervisor Prof. Patrick Cramer for giving me the opportunity to work in such an inspiring atmosphere and for giving me the freedom to pursue a variety of exciting biological questions and collaborations over the years. His way of running a laboratory and leading people will hopefully influence my future professional and personal life. Thank you, Patrick, for the trust and the personal support during those exciting four years!

I am also very thankful to the members of my Thesis Advisory Committee, Prof. Dietmar Martin and Dr. Heidi Feldmann, for their advice, constant support, help with “yeast issues” and numerous discussions. At the same time I want to thank Steve Hahn, an inspiring researcher and certainly one of the nicest persons I have ever met. Thanks Steve, for letting me stay in your lab, for teaching me that a good biochemical experiment needs more controls than actual samples, for writing reference letters and for many great discussions in the lab, in restaurants and at conferences.

Next, I want to thank a great scientist, curious investigator, long-standing collaborator and very dear friend. “Merci beaucoup” Laurent, for sharing projects, first-authorships, ideas, lunch/coffee breaks, homemade “Foie gras with Périgord truffle” and loads of chocolate with me. You taught me many critical methods including protein purification, crystallization, data processing and preparing the world’s most delicious brownie.

At the same time, I want to thank you, Larissa, for being a great colleague and very good friend over the past years. Without you taking over a huge amount of the experimental workload and keeping me company during all those lunch/coffee breaks, many projects would not have been successful. I am also very thankful to Rike, Erika and Sonja, my present and former “bench neighbors” and good friends, for proofreading as well as sharing ideas, projects, jokes, lunch/coffee breaks, buffers, protocols, hiking trips...

Having many collaborations of course means many more people to thank: Fabian and Toni, both outstanding students, for their dedication to the projects and their excellent scientific contributions; Claudia Buchen, Stefan Benkert, Kristin Leike, Nicole Pirkl and Steffi Etzold for lots of help, for keeping the laboratory running and for always trying to make it a better place to work at; Elmar and Tobias for discussions, sharing projects, bike trips to retreats; “The Array Group” including Andi, Daniel, Kerstin and Mai for numerous late-night discussions, critical comments, help with ChIP, gene expression profiling, DTA and data analysis; Jasmin for lots of fun at the CSH meeting and for letting me win all those Squash matches; and of course all former and present members of the Cramer laboratory for lots of help and the great atmosphere.

Furthermore, I want to thank Mirijam Zeller from the Thomm lab for an exciting collaboration and lots of fun; Axel Imhof and Ignasi Forne for the collaboration on label-free tandem mass spectrometry; Johannes Soeding and Holger Hartmann for many discussions and sharing projects; Lina, Sittinan and Britta from the Straesser lab for help with yeast problems and for always helping out with buffers, antibodies, plates and great ideas; James, Neeman, Linda and Hung-Ta from the Hahn lab for lots of help and discussions, for being good friends and for providing a place to stay whenever I am in the US; Kai Hell and all former and present members of the ENB graduate school “Protein dynamics in health and disease” for excellent scientific talks, discussions and legendary retreats; the Elite network of Bavaria for funding.

I am also very grateful to the Boehringer Ingelheim Fonds for granting me one of their PhD fellowships. I want to particularly thank Claudia, Monika and all the people working at the BIF for the continuous personal and professional support. I am really happy to be a member of the BIF family!

Liebe Mama, lieber Papa und liebe Christina, vielen vielen Dank für eure Unterstützung über all die Jahre! Diese Doktorarbeit wäre ohne Euch sicher nicht möglich gewesen!

Liebe Kati, ich danke dir für all die schönen gemeinsamen Jahre. Du hast mir immer Halt und ein normales Leben gegeben, selbst in den stressigsten Zeiten.

## Summary

In eukaryotes, transcription of protein-coding genes and many non-coding RNAs relies on RNA polymerase II, a common set of general transcription factors (GTF), namely TFIIA, -B, -D, -E, -F, -H, -S, and coactivator complexes such as Mediator. The GTF are required for core promoter recognition, promoter opening, transcription start site recognition and initial RNA synthesis. The large multiprotein complex Mediator promotes preinitiation complex (PIC) formation at the core promoter by bridging gene-specific activators bound to upstream DNA elements to the basal transcription machinery. The 7-subunit Mediator head module plays a pivotal role during PIC formation, contacting Pol II-TFIIF, TATA-binding protein (TBP), TFIIB, and TFIH.

During this work a combination of X-ray crystallography, yeast genetics, biochemical assays, chromatin immunoprecipitation and genome-wide expression profiling was used to elucidate the submodular architecture and molecular mechanisms underlying Mediator head function. The conserved functional head submodules Med8C/18/20 and Med11/22 were identified and their distinct functions characterized. A conserved flexible anchoring mode to the head module was demonstrated for both submodules. While the non-essential Med8C/18/20 is required for low transcription levels of a specific subset of nonactivated genes, mutations in the essential Med11/22 submodule had a more pleiotropic effect on gene expression. Structure-guided mutagenesis of Med11/22 identified a highly conserved surface patch required for stable PIC formation *in vitro* and *in vivo*. Furthermore, the determined crystal structure of Med11/22 revealed an unexpected homology to the Med7/21 middle module subcomplex. Structure predictions identified a total of 9 out of 17 Mediator core subunits and two metazoan-specific subunits sharing the heterodimeric four-helix bundle fold of Med11/22. During evolution this common structural building block appears to have duplicated and diversified to generate new protein interaction surfaces and thus accommodate the need for more complex regulatory mechanisms.

Furthermore, reliable nuclear extract based assays for *in vitro* transcription and PIC assembly on native yeast promoter templates were established and optimized. The assays were used in several collaborations to characterize the role of individual factors/subcomplexes during Pol II transcription. In addition, a label-free mass spectrometry approach was used to identify factors depending on the Mediator head module for core promoter binding.

## Publications

Parts of this work have been published or are in the process of publication. :

Lariviere L<sup>1</sup>, Seizl M<sup>1</sup>, van Wageningen S, Roether S, Feldmann H, Straesser K, Hahn S, Holstege F, Cramer P. Structure-system correlation identifies a gene regulatory Mediator submodule. *Genes Dev.* 2008 Apr 1;22(7):872-877

<sup>1</sup> these authors contributed equally

Koschubs T, Seizl M, Larivière L, Kurth F, Baumli S, Martin DE, Cramer P. Identification, structure, and functional requirement of the Mediator submodule Med7N/31. *EMBO J.* 2009 Jan 7;28(1):69-80. Epub 2008 Dec 4

Kostrewa D<sup>1</sup>, Zeller ME<sup>1</sup>, Armache KJ<sup>1</sup>, Seizl M, Leike K, Thomm M, Cramer P. RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature.* 2009 Nov 19;462(7271):323-30

<sup>1</sup> these authors contributed equally

Vojnić E, Mourão A, Seizl M, Simon B, Wenzek L, Larivière L, Baumli S, Baumgart K, Meisterernst M, Sattler M, Cramer P. The Mediator Med25 activator interaction domain: Structure and cooperative binding of VP16 subdomains. *Nat Struct Mol Biol.* 2011 Apr;18(4):404-9. Epub 2011 Mar 6

Czeko E, Seizl M, Mielke T, Cramer P. Iwr1 directs RNA polymerase II nuclear import. *Mol Cell.* 2011 in press

Chanarat S, Seizl M, Strässer K. The Prp19 complex is a novel transcription elongation factor. *Genes Dev.* 2011 in press

Seizl M<sup>1</sup>, Lariviere L<sup>1</sup>, Pfaffeneder T, Wenzek L, Cramer P. Mediator head subcomplex Med11/22 contains a common helix bundle building block with a specific function in transcription initiation complex stabilization. *Nucleic Acids Res.* 2011 in press

<sup>1</sup> these authors contributed equally

## Table of contents

Erklärung	I
Ehrenwörtliche Versicherung	I
Acknowledgement	II
Summary	IV
Publications	V
Table of contents	VI
<b>1 Introduction</b>	<b>1</b>
<b>1.1 RNA polymerases</b>	<b>1</b>
<b>1.2 The RNA polymerase II transcription cycle</b>	<b>3</b>
1.2.1 Initiation and the general transcription factors	3
1.2.2 Elongation and the CTD-code	5
1.2.3 Termination and reinitiation	6
<b>1.3 Regulation of eukaryotic transcription initiation</b>	<b>6</b>
1.3.1 Core promoter elements	6
1.3.2 Promoter accessibility and chromatin	7
1.3.3 Positive cofactors	9
1.3.4 Negative cofactors	11
<b>1.4 The general coactivator complex Mediator</b>	<b>11</b>
1.4.1 Discovery and conservation	11
1.4.2 Modular architecture and structure	13
1.4.3 Mediator function in transcriptional regulation	13
1.4.4 Head module architecture and function	14
<b>1.5 Aims and scope of this work</b>	<b>16</b>
<b>1.6 Additional contributions</b>	<b>17</b>
<b>2 Materials and Methods</b>	<b>18</b>
<b>2.1 Materials</b>	<b>18</b>
2.1.1 Strains	18
2.1.2 Plasmids	19
2.1.3 Oligonucleotides	25
2.1.4 Antibodies	26
2.1.5 Growth media and additives	26
2.1.6 Buffers and solutions	27

<b>2.2</b>	<b>Methods</b> .....	<b>30</b>
2.2.1	Molecular cloning .....	30
2.2.2	General protein methods.....	31
2.2.3	Recombinant protein purification protocols .....	36
2.2.4	<i>In vitro</i> biochemical assays .....	40
2.2.5	Yeast methods .....	43
2.2.6	Gene-expression profiling .....	46
<b>3</b>	<b>Results and Discussion</b> .....	<b>49</b>
<b>3.1</b>	<b>Structure-system correlation defines the gene regulatory Mediator submodule Med8C/18/20</b> .....	<b>49</b>
3.1.1	Med8C/18/20 is a subcomplex of the Mediator head .....	49
3.1.2	Conserved Med8C/18/20-core head interface .....	51
3.1.3	Med8C/18/20 is required for activated transcription <i>in vitro</i> .....	52
3.1.4	Med8C/18/20 is a functional submodule <i>in vivo</i> .....	53
3.1.5	Med8C/18/20 is required for transcription of specific genes .....	55
3.1.6	Down-regulation of nonactivated genes and basal promoter activity.....	55
3.1.7	Discussion .....	56
<b>3.2</b>	<b>Mediator head subcomplex Med11/22 contains a common helix bundle building block with a specific function in transcription initiation complex stabilization</b> .....	<b>58</b>
3.2.1	Revised Med11 N-terminus.....	58
3.2.2	Med11/22 structure solution .....	58
3.2.3	A helix bundle building block in Mediator .....	61
3.2.4	Essential C-terminal helices extend from the Med11/22 helix bundle.....	62
3.2.5	Med11/22 extensions bind a Med17 C-terminal domain .....	63
3.2.6	A conserved Med17C/11/22 Mediator subcomplex.....	65
3.2.7	A highly conserved interaction patch on Med11/22.....	65
3.2.8	Med11/22 is a functionally distinct submodule .....	68
3.2.9	The Med11/22 surface patch functions in PIC stabilization.....	69
3.2.10	Discussion .....	70
<b>3.3</b>	<b>Mediator head module controls preinitiation complex formation</b> .....	<b>72</b>
3.3.1	A recombinant head module is functionally active .....	72
3.3.2	Preinitiation complex formation requires the head module .....	74
3.3.3	A label-free mass spectrometry approach to study PIC assembly.....	76

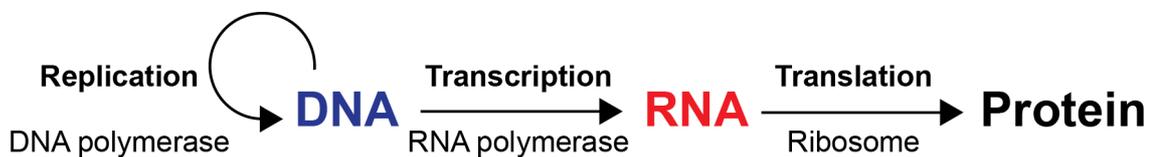
---

3.3.4	Mediator body is recruited independent of head module .....	78
3.3.5	Head module is only required for the recruitment of the basal machinery .	80
3.3.6	Discussion .....	80
<b>3.4</b>	<b>Additional contributions.....</b>	<b>82</b>
3.4.1	Overview .....	82
3.4.2	Mediator middle submodule Med7N/31 cooperates with TFIIIS during activated transcription <i>in vitro</i> .....	82
3.4.3	The archetypical activator VP16 targets the Mediator subunit Med25 through a conserved synergistic use of subdomains.....	86
3.4.4	General transcription factor TFIIIB controls transcription start site selection.....	90
<b>4</b>	<b>Conclusion and Outlook .....</b>	<b>93</b>
	<b>Abbreviations .....</b>	<b>107</b>
	<b>List of Figures.....</b>	<b>109</b>
	<b>List of tables .....</b>	<b>110</b>
	<b>Curriculum vitae .....</b>	<b>111</b>

# 1 Introduction

## 1.1 RNA polymerases

The central dogma of molecular biology was first postulated by Francis Crick (Crick, 1970). It describes the directional flow of genetic information from deoxyribonucleic acid (DNA) via ribonucleic acid (RNA) to proteins (Figure 1). The underlying fundamental biological processes of DNA replication, transcription and translation are found in all three kingdoms of life.



**Figure 1: The central dogma of of molecular biology**

Transcription, the process of synthesizing RNA from a DNA template, is carried out by DNA-dependent RNA polymerases (RNAP). In 1960, four laboratories had simultaneously discovered RNAP in *E. coli* lysate (Hurwitz et al, 1960; Stevens, 1960), rat liver nuclei (Weiss & Gladstone, 1959) and extracts from pea (Huang et al, 1960). In general, RNAPs can be divided into single-subunit and multisubunit enzyme families. The two families lack any sequence or structural homology and are probably the product of convergent evolution (Cramer, 2002a). Single-subunit RNAP are found in bacteriophages (e.g. T3 and T7) and mitochondria while multi-subunit RNA polymerases are found in all three kingdoms of life. Whereas bacteria and archaea rely on a single RNAP for RNA synthesis, eukaryotic cells contain at least three distinct enzymes (Roeder & Rutter, 1969). RNA polymerase I (Pol I) is located in the nucleoli and synthesizes 5.8S, 18S and 28S ribosomal RNA (rRNA). RNA Polymerase II (Pol II) is located in the nucleoplasm and synthesizes the messenger RNA (mRNA) of all protein coding genes, small nucleolar RNAs (snoRNAs) and some small nuclear RNAs (snRNAs). Pol III is also located in the nucleoplasm and synthesizes transfer RNAs (tRNAs), 5S rRNA, some snRNAs and other small RNAs. The recently discovered

plant-specific Pol IV and Pol V are dedicated to the formation and maintenance of heterochromatin through small interfering RNAs (siRNAs) (Lahmy et al, 2010).

Although multisubunit RNAP differ widely in their subunit composition and molecular weight (Table 1), the structural core and enzymatic mechanism is conserved in all three kingdoms of life (Cramer, 2002b; Hirata & Murakami, 2009; Ream et al, 2009).

**Table 1: Subunit composition of multisubunit RNAP from all three kingdoms of life**

	Pol I	Pol II	Pol III	Pol IV	Pol V	Archaea	Bacteria
Core	A190	Rpb1	C160	Nrpd1	Nrpe1	A' + A''	β'
	A135	Rpb2	C128	At3g23780	At3g23780	B' + B''	β
	AC40	<b>Rpb3</b>	AC40	<b>Rpb3</b>	Rpb3/ Nrpe3b	D	α
	AC19	<b>Rpb11</b>	AC19	<b>Rpb11</b>	<b>Rpb11</b>	L	α
	<b>Rpb6</b>	<b>Rpb6</b>	<b>Rpb6</b>	<b>Rpb6</b>	<b>Rpb6</b>	K	ω
	<b>Rpb5</b>	<b>Rpb5</b>	<b>Rpb5</b>	<b>Rpb5</b>	Nrpe5	H	-
	<b>Rpb8</b>	<b>Rpb8</b>	<b>Rpb8</b>	<b>Rpb8</b>	<b>Rpb8</b>	-	-
	<b>Rpb10</b>	<b>Rpb10</b>	<b>Rpb10</b>	<b>Rpb10</b>	<b>Rpb10</b>	N	-
	<b>Rpb12</b>	<b>Rpb12</b>	<b>Rpb12</b>	<b>Rpb12</b>	<b>Rpb12</b>	P	-
	A12.2	<b>Rpb9</b>	C11	<b>Rpb9</b>	<b>Rpb9</b>	-	-
Stalk	A14	Rpb4	C17	At4g15950	At4g15950	F	-
	A43	Rpb7	C25	Nrpd7	Nrpe7	E	-
TFIIF-like	A49	(Tfg1) <sup>b</sup>	C37	?	?	-	-
	A34.5	(Tfg2) <sup>b</sup>	C53	?	?	-	-
Pol III-specific	-	-	C82	-	-	-	-
	-	-	C34	-	-	-	-
	-	-	C31	-	-	-	-
Subunits	14	12	17	12	12	12	5

table adapted from (Cramer, 2002b; Hirata & Murakami, 2009; Ream et al, 2009)

<sup>a</sup> Subunits shared between Pol II and other RNAP are highlighted in bold red

<sup>b</sup> (Geiger et al, 2010)

Differences in peripheral subunits of the RNAP and regulatory factors reflect the complexity of the respective system. In bacteria, RNAP initiation relies on a single regulatory factor sigma, which recognizes promoter sequences and recruits the enzyme directly (Mooney et al, 2005). In archaea, two factors, TFB and the TATA-binding protein (TBP), are essential for transcription initiation (Geiduschek & Ouhammouch, 2005). The much bigger eukaryotic RNA polymerases depend on large sets of regulatory factors that differ between Pol I, Pol II and Pol III (Roeder, 1996).

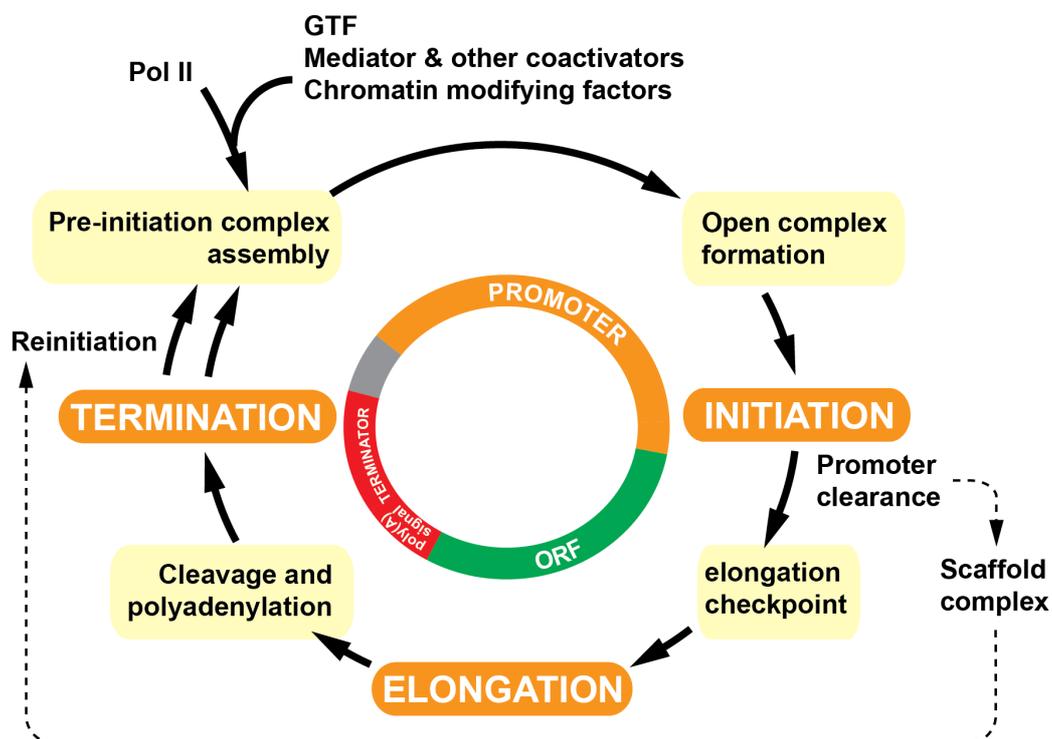
The highly regulated transcription of protein-coding genes by Pol II relies on a plethora of factors, including general transcription factors (GTF), negative and positive cofactors including chromatin modifying factors and coactivator complexes, elongation factors, RNA processing factors and termination factors (Hahn, 2004; Svejstrup, 2004). The unique C-terminal domain (CTD) of the largest subunit of Pol II serves a flexible binding

platform for many of these regulatory factors. The CTD consists of multiple highly conserved heptapeptide tandem repeats of the sequence Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7 (Allison et al, 1985; Corden et al, 1985), which are heavily phosphorylated during transcription (Cadena & Dahmus, 1987). CTD truncations to less than 8 heptarepeats or mutations of the phosphorylation sites Tyr1, Ser2 or Ser5 are lethal in yeast (West & Corden, 1995). Changes in the CTD phosphorylation pattern were shown to orchestrate the association of different sets of regulatory factors required during different phases of transcription (Phatnani & Greenleaf, 2006).

## 1.2 The RNA polymerase II transcription cycle

### 1.2.1 Initiation and the general transcription factors

Eukaryotic transcription by Pol II has been structurally and biochemically characterized in detail.

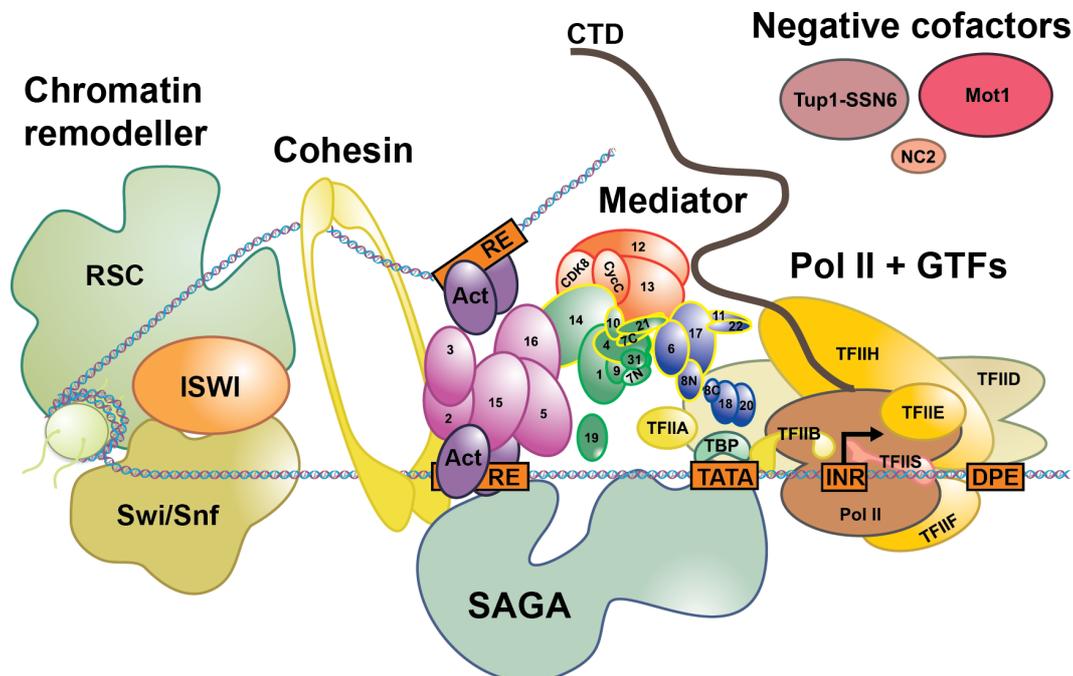


**Figure 2: RNA polymerase II transcription cycle**

The main phases of the transcription cycle are colored orange, important events of regulation are highlighted in yellow. The circle in the middle depicts the occurrence of the events in relation to the position on the gene. GTFs = general transcription factors; ORF = open reading frame (Image courtesy of Stefan Dengl, Gene Center Munich). Adapted from (Hahn, 2004; Svejstrup, 2004)

Pol II transcription follows the so-called transcription cycle (Figure 2), which is divided into initiation, elongation, termination, and reinitiation (Hahn, 2004; Svejstrup, 2004). Each phase is regulated and requires a specific set of factors.

Transcription initiation is a multi-step process involving over a 100 different polypeptides (Sikorski & Buratowski, 2009; Thomas & Chiang, 2006; Venters et al, 2011). First, gene-specific activator proteins (activators) recognize and bind their cognate DNA motif, which is often located within the upstream-activating sequence (UAS). Activators can bind to the UAS of multiple genes to generate a highly coordinated regulation of gene expression. However, often a gene is controlled by the action of multiple activators, leading to combinatorial regulation. Next, chromatin remodelers/modifiers and coactivators, such as ISWI, RSC, SWI/SNF, NuA4, RSC, INO80, SAGA and Mediator, are recruited to alter the chromatin environment and promote preinitiation complex (PIC) formation at the core promoter (Figure 3) The PIC comprises Pol II and the GTFs, namely TFIIA, -B, -D, -E, -F, -H and -S. The GTFs are required for core promoter recognition, promoter opening, transcription start site selection, and initial RNA synthesis (Table 2).



**Figure 3: Regulatory factors and the basal transcription machinery at the core promoter**

Complexes involved in Pol II transcription initiation and important cis-regulatory elements are depicted (RE = regulatory element, TATA = TATA box, INR = Initiator, DPE = downstream promoter element). The transcription start site is indicated with an arrow. Gene looping through Mediator-Cohesin interaction is indicated. Modular architecture of the coactivator complex Mediator is shown (head = blue, middle = green, tail = pink, kinase = orange). Mediator subunits essential for yeast viability are outlined in yellow. All proteins and complexes are drawn approximately at relative scale. Image courtesy of T. Koschubs, Gene Center.

Upon ATP hydrolysis the promoter DNA is melted and the template strand is positioned into the active cleft of Pol II (open complex formation) (Wang et al, 1992). With the help of TFIIB, the transcription start site (TSS) is recognized and RNA synthesis is initiated (initiation) (Cho & Buratowski, 1999; Pardee et al, 1998; Ranish et al, 1999). When the transcript reaches a length of at least seven nucleotides TFIIB is released and Pol II enters the elongation phase (promoter clearance) (Pal et al, 2005). Most of the initiation factors are left behind in the promoter-bound scaffold complex, which allows rapid transcription reinitiation of previously transcribed genes (Yudkovsky et al, 2000).

### **1.2.2 Elongation and the CTD-code**

The elongating Pol II transcription complex undergoes several transitions accompanied by changes in the phosphorylation pattern of the CTD (CTD code). The CTD code determines the factors associating with Pol II ensuring cotranscriptional RNA processing and chromatin modifications (Buratowski, 2009; Meinhart et al, 2005; Phatnani & Greenleaf, 2006; Proudfoot et al, 2002). Upon transcription initiation the CTD of Pol II starts getting phosphorylated at Ser5 and Ser7 by the CTD kinases Kin28 (TFIIH) and Cdk8 (Mediator) facilitating promoter escape (Liu et al, 2004). Shortly after transcription initiation, the heterodimeric elongation factor Spt4/5 (DSIF) associates with Pol II (Wada et al, 1998). In higher eukaryotes, the negative elongation factor (NELF) traps the transcription complex at promoter proximal sites (Yamaguchi et al, 1999). The capping enzymes are recruited and the nascent mRNA is capped (Wen & Shatkin, 1999). The CTD kinase P-TEFb subsequently phosphorylates the C-terminal region of Spt5 and the CTD at Ser2 causing a release of the transcription machinery and the entry into productive transcript elongation (Kim & Sharp, 2001). In yeast, which lacks the mechanism of promoter proximal stalling, all elongation factors are recruited in a single 5' transition approximately 150 nt downstream of the transcription start site (Mayer et al, 2010). During elongation, the complex interplay of CTD phosphatases and the CTD kinases Ctk1 and Bur1 changes the CTD code from high Ser5/Ser7 phosphorylation at the 5' end of the gene to high Ser2 phosphorylation at the 3' end of the gene (Buratowski, 2009).

### **1.2.3 Termination and reinitiation**

When the transcription machinery encounters the poly (A) site at the 3' end of genes the final phase of the transcription cycle is reached and the elongation factors are replaced by termination factors (Richard & Manley, 2009). During this phase transcript cleavage, polyadenylation and termination occurs. Similar to the preceding phases the underlying processes are highly regulated and require a specific set of factors, such as Pcf11, Rat1 and Rai1. At protein-coding genes, termination is coupled to RNA processing events and often occurs far downstream of the poly (A) site (Buratowski, 2005; Proudfoot, 1989). The underlying mechanisms are still poorly understood.

After termination, Pol II is released from the template DNA and can enter facilitated reinitiation through association with the promoter bound scaffold complex (Yudkovsky et al, 2000). This scaffold complex comprises TFIIA, -D, -E, -H and Mediator, and is stabilized by gene-specific activators. The direct recruitment of Pol II into the preformed complex was suggested to enable rapid PIC formation and consequently increased levels of gene transcription.

## **1.3 Regulation of eukaryotic transcription initiation**

### **1.3.1 Core promoter elements**

Core promoter elements (CPE) are conserved cis-regulatory DNA motifs required for promoter recognition, PIC stability, promoter opening and transcription initiation. In eukaryotes, at least seven different CPE have been identified so far (Juven-Gershon et al, 2008). CPE are not universally present in all promoters, but occur in different combinations thus adding an additional layer of regulation (Müller et al, 2007).

The evolutionary most ancient CPE, which are found in archaeal and eukaryotic promoters, comprise the TATA box, the TFIIB-recognition element (BRE) and the initiator element (INR) (Soppa, 1999). The TATA box was the first described CPE and comprises the highly conserved consensus motif TATAAWWR. In most eukaryotes, it is located 25-30 bp upstream of the transcription start site (TSS) and marks the assembly point for the PIC. In yeast, the distance is more variable (40-120 bp) and therefore Pol II has to scan for the TSS after open complex formation (Struhl, 1989). In all species, the TATA box binding protein (TBP) binds directly to the TATA box bending

the promoter DNA (Kim et al, 1993a; Kim et al, 1993b). The TATA box is flanked up- and downstream by BREs, the BRE<sup>u</sup> and BRE<sup>d</sup>, respectively (Deng & Roberts, 2006). TFIIB, a GTF directly interacting with TBP and Pol II, contacts these elements further stabilizing the PIC and enforcing directionality (Tsai & Sigler, 2000). The INR overlaps with the TSS and is sufficient for directing accurate transcription from TATA-less promoters through direct interaction with TFIID (Chalkley & Verrijzer, 1999; Smale, 1997). Interestingly, promoters of trichomonads (a very early-diverging eukaryotic lineage) seem to completely lack functional TATA boxes and rely on a highly conserved INR instead (Liston & Johnson, 1999).

Other CPE described in metazoan promoters, like the downstream core promoter element (DPE), the motif ten element (MTE) and the downstream core element (DCE), are even less understood and further complicate the underlying regulatory mechanisms (Juven-Gershon et al, 2008). All efforts to identify CPE, other than the TATA box and a degenerate INR, in the well-characterized yeast species *Saccharomyces cerevisiae* (Sc) failed so far. Nevertheless, recent genome-wide studies revealed that TATA boxes are found in less than 20% of yeast and metazoan promoters, clearly indicating the presence of additional CPE (Basehoar et al, 2004). Identification of these elements using bioinformatics might be particularly difficult because they are either very degenerate, relying more on biophysical properties than DNA sequence or because other determinants like chromatin environment play an important role.

### **1.3.2 Promoter accessibility and chromatin**

In contrast to the “naked” DNA in bacteria, the genomic DNA in eukaryotes, and to some extent in archaea, is bound by histone proteins (Reeve, 2003; Sandman & Reeve, 2005). Therefore, most CPE are not readily accessible for binding. In eukaryotic cells, a 146 base pair DNA stretch is wrapped around a histone octamer, formed by two H2A/H2B and two H3/H4 heterodimers (Luger et al, 1997). The so-called nucleosomes represent the primary unit of chromatin structure. Chromatin serves a dual role within the cell by compacting the genome and by restricting the access of DNA-binding factors (Li et al, 2007).

In bacteria, where DNA is always accessible, the RNAP has an inherent, sigma factor-directed promoter-binding ability (Mooney et al, 2005). Therefore, binding affinities represent the major determinant of promoter strength and regulatory factors merely increase the affinity of RNAP.

In contrast, eukaryotic DNA is generally inaccessible and Pol II does not have inherent promoter binding abilities but relies on a set of positive cofactors, namely chromatin remodelers, chromatin modifiers and coactivator complexes, to promote PIC formation and transcription initiation (Table 2) (Sikorski & Buratowski, 2009; Thomas & Chiang, 2006).

**Table 2: Factors involved in Pol II transcription initiation**

Factor	Subunits	Function
Pol II	12	synthesizes all mRNAs and a subset of non-coding RNAs including snoRNAs and snRNAs
TFIIA	2	stabilizes TBP at the TATA box; can act as a coactivator by interacting with activators; can counteract repressive effects of negative cofactors like NC2
TFIIB	2	required for accurate TSS selection; interacts with Pol II and TBP stabilizing the PIC; stabilizes binding of TFIID to the promoter;
TFIID	15	comprises TBP and TBP-associated factors (TAFs); nucleates PIC assembly either through TBP binding to TATA sequences or TAF binding to other CPE; coactivator activity through direct interaction of TAFs and gene specific activators
TFIIE	2	helps recruit TFIIH to promoters; stimulates helicase and kinase activities of TFIIH; binds ssDNA and is essential for promoter melting
TFIIF	2-3	tightly associates with Pol II; enhances affinity of Pol II for TBP-TFIIB-promoter complex; necessary for recruitment of TFIIE/TFIIH to the PIC; aids in start site selection and promoter escape; enhances elongation efficiency
TFIIH	10	ATPase/helicase necessary for promoter opening and promoter clearance; helicase activity for transcription coupled DNA repair; kinase activity required for phosphorylation of Pol II CTD; facilitates transition from initiation to elongation
TFIIS	1	stimulates intrinsic transcript cleavage activity of Pol II allowing backtracking to resume RNA synthesis after transcription arrest; stimulates PIC assembly at some promoters
Mediator	25	bridges interaction between activators and basal factors; stimulates both activator dependent and basal transcription; required for transcription from most Pol II dependent promoters
SAGA	20	interacts with activators, histone H3, and TBP; histone acetyltransferase activity; deubiquitinating activity
Tup1-Ssn6	2	represses transcription initiation through repressive nucleosome positioning, histone deacetylation and interference with Mediator function
Mot1	1	induces dissociation of TBP/DNA complexes in ATP dependent manner; can have both positive and negative effects on transcription
NC2	2	binds TBP/DNA complexes and blocks PIC assembly; can have both positive and negative effects on transcription

adapted from (Sikorski & Buratowski, 2009; Thomas & Chiang, 2006)

Based on recent genome-wide nucleosome occupancy studies, promoters in eukaryotes can be classified into two contrasting categories, “open” and “covered”, comprising constitutive and highly regulated genes, respectively (Cairns, 2009; Field et al, 2008; Tirosch & Barkai, 2008). Computational analysis and predictions of nucleosome positions revealed that promoter accessibility is highly dependent on biophysical properties, like DNA curvature or stiffness (Segal & Widom, 2009; Struhl, 1985). Open promoters contain an approximately 150 bp nucleosome-depleted region

directly upstream of the TSS enabling direct access for transcription factor binding. A combination of two sequence elements into a tripartite structure stabilizes this open architecture: a central poly(dA:dT) tract flanked by nucleosome positioning sequences, fixing the so-called -1 and +1 nucleosome in place (Segal et al, 2006; Struhl, 1985; Yuan et al, 2005). Interestingly, most of the open promoters lack a functional TATA box and contain the H2A variant H2A.Z (htz1 in yeast) at the +1 or -1 nucleosome (Raisner et al, 2005; Zhang et al, 2005). In contrast, in covered promoters TATA boxes are more enriched, but all CPE and most of the other cis-regulatory elements are masked through nucleosome binding (Basehoar et al, 2004; Huisinga & Pugh, 2004). Often a single transcription factor binding site is exposed, which is bound upon activation by a “pioneering” transcription factor (Cairns, 2009). Subsequent recruitment of chromatin remodelers and modifiers makes the core promoter accessible and enables recruitment of coactivator complexes facilitating PIC formation (Becker & Hörz, 2002; Sikorski & Buratowski, 2009). Therefore, chromatin structure and consequently chromatin remodeling represents an important additional layer of regulation in eukaryotic gene regulation.

### **1.3.3 Positive cofactors**

Positive cofactors are typically subdivided based on their mechanism of facilitating Pol II transcription initiation into two classes, namely coactivators and chromatin remodelers/modifiers (Sikorski & Buratowski, 2009; Thomas & Chiang, 2006). However, many complexes have multiple roles making a clear classification difficult.

Multiprotein coactivator complexes, such as the Mediator complex (Malik & Roeder, 2010), the Spt-Ada-Gcn5-acetyltransferase complex (SAGA) (Rodríguez-Navarro, 2009) and the basal transcription factor TFIID (Cler et al, 2009), are highly conserved from yeast to man. Gene specific activators recruit these coactivators to the core promoter, where they facilitate and stabilize PIC formation through multiple contacts with the basal transcription machinery (Sikorski & Buratowski, 2009). Whereas Mediator is considered a general coactivator complex required for transcription of most Pol II genes (see Chapter 1.4) (Kornberg, 2005), the role of SAGA and TFIID seems to be partially redundant (Basehoar et al, 2004; Huisinga & Pugh, 2004; Lee et al, 2000). While some activators, like the acidic activators Gal4 and Gcn4 target both complexes (Reeves & Hahn, 2005), other activators, like Rap1 seem to recruit only TFIID (Mencía et al, 2002). Recent genome-wide chromatin immunoprecipitation (ChIP) studies in

yeast revealed that highly regulated TATA-containing genes rely more on SAGA while constitutive TATA-less genes rely more on TFIID (Basehoar et al, 2004; Huisinga & Pugh, 2004). This is in accordance with previous observations in higher eukaryotes, where TFIID was shown to recognize and bind directly to the INR, not requiring a functional TATA box for transcription activation (Chalkley & Verrijzer, 1999). Taken together, several modes of promoter recognition and PIC formation seem to exist, which are determined by activator-coactivator and CPE-coactivator interactions (Bhaumik, 2011; Sikorski & Buratowski, 2009).

SAGA and TFIID appear similar in shape and share several TBP-associated factors (TAFs). Interestingly, many TAFs contain a conserved histone fold domain and form heterodimeric subcomplexes. These heterodimers appear to be a common building block important for the architecture and function of both complexes (Cler et al, 2009; Leurent et al, 2002). Nevertheless, the two complexes differ in several key features, which helps to explain their mechanistic specificity. The multisubunit TFIID comprises the TATA-binding protein (TBP) and 13-14 different TAFs, some of which appear to be present in at least two copies (Cler et al, 2009). The TFIID-specific subcomplex Taf1/Taf2 was shown to recognize and bind the INR (Kaufmann & Smale, 1994). In addition, the TFIID-specific subunit Taf3 interacts through a plant homeodomain (PHD) finger with trimethylated histone H3 at lysine 4 (H3K4me3) facilitating recruitment of TFIID. The recruitment is further stimulated by acetylation of H3 at lysine 9 (H3K9) and lysine 14 (H3K14) (Taverna et al, 2007).

The multiprotein complex SAGA comprises 21 subunits in yeast. Only the 5 TAFs shared with TFIID and the ATM/PI-3-kinase Tra1 shared with NuA4 complex are essential for yeast viability (Bhaumik, 2011; Rodríguez-Navarro, 2009). In contrast to TFIID, TBP is not a stable subunit of SAGA, but interacts transiently with the SAGA-specific subunit Spt3 during transcription activation (Mohibullah & Hahn, 2008). The two SAGA-specific subunits Gcn5 and Ubp8 possess histone acetyltransferase and deubiquitylase activities, respectively. Both subunits were shown to regulate gene expression of a subset of genes through modulation of the chromatin structure (Henry et al, 2003).

### 1.3.4 Negative cofactors

In addition to the plethora of positive cofactors, three negative cofactors have been described in yeast so far (Sikorski & Buratowski, 2009). The Ssn6-Tup1 complex is considered a global repressor complex, antagonizing transcriptional activation through repressive nucleosome positioning, histone deacetylation and interference with Mediator function (Zhang & Reese, 2004). Interestingly, the other two negative cofactors, Mot1 and the heterodimeric NC2 complex, were identified in an extragenic suppressor screen of a temperature-sensitive Mediator mutation (Gadbois et al, 1997; Lee et al, 1998). Mutations in these factors compensate the loss of Mediator function at elevated temperatures. Both cofactors act very specifically through direct interactions with TBP. Mot1, a Snf2 ATPase, removes TBP from the promoter, whereas NC2 blocks TBP-TFIIA and TBP-TFIIB interactions (Auble, 2009; Pereira et al, 2003). While the mode of action is very different, the outcome of interfering with PIC formation and stability is the same. Interestingly, recent genome-wide CHIP studies found NC2 and Mot1 at many active yeast promoters (van Werven et al, 2008). While NC2 and Mot1 repress TATA box containing promoters, they seem to activate TATA-less promoters. Although the underlying mechanism is still unclear, it was suggested that Mot1 and NC2 might remove non-functional TBP from the promoter to enable binding of other factors (Sikorski & Buratowski, 2009).

## 1.4 The general coactivator complex Mediator

### 1.4.1 Discovery and conservation

First evidence for the existence of Mediator came from squelching experiments in yeast (Gill & Ptashne, 1988) and mammalian cells (Triezenberg et al, 1988). Overexpression of one activator interfered with the activation of Pol II transcription by another activator. At the beginning the effect was attributed to competitive binding to the same target within the basal transcription machinery. However, addition of an excess of Pol II or any GTF to a crude yeast *in vitro* transcription system did not relieve the squelching effect, while addition of a crude yeast extract did (Flanagan et al, 1991; Kelleher et al, 1990). This strongly indicated the existence of an intermediate layer of regulation between the activators and the transcription machinery in eukaryotes. The unknown factor was termed Mediator, but it took another four years until the Mediator complex

was first purified and characterized (Kim et al, 1994). In parallel, genetic screens in yeast had identified several suppressors of the cold-sensitive phenotype of a Pol II CTD truncation. The identified extragenic suppressors were termed suppressors of RNA polymerase B (Srb) (Nonet & Young, 1989). Later on, all 9 Srb proteins turned out to be subunits of the large coactivator complex Mediator.

Mediator was subsequently purified from various fungi, metazoans and a plant (Bäckström et al, 2007; Boube et al, 2002). Comparative genomics identified an ancient 17-subunit core Mediator, which is conserved in all eukaryotes (Bourbon, 2008) (Table 3).

**Table 3: Mediator subunit composition and modular architecture**

Module	Subunit <sup>a</sup>	Sc	Hs	pdb entry <sup>b</sup>
Head	<u>Med6</u>	Med6	hMed6/Drip33	-
	<u>Med8</u>	Med8	Arc32	3COT/2HZS
	<u>Med11</u>	Med11	HSPC296	this study
	<u>Med17</u>	Srb4	TRAP80/DRIP77/CRSP77	-
	<u>Med18</u>	Srb5	p28b	3COT/2HZM/2HZS
	<u>Med20</u>	Srb2	hTRFP/p28a	2HZM/2HZS
	<u>Med22</u>	Srb6	Med24/Surf5	this study
Head/Middle	Med19	Rox3	LCMR1	-
Middle	Med1	Med1	TRAP220	-
	<u>Med4</u>	MED4	TRAP36/DRIP36	-
	<u>Med7</u>	Med7	hMed7/Drip34	3FBI/3FBN/1YKE/1YKH
	<u>Med9</u>	Med9/Cse2	Med25	-
	<u>Med10</u>	Med10/Nut2	hMed10/hNut2	-
	<u>Med21</u>	Srb7	hSrb7/p21	1YKE/1YKH
	<u>Med31</u>	Soh1	hSoh1	3FBI/3FBN
Middle/Tail	<u>Med14</u>	Rgr1	TRAP170/DRIP150/CRSP150	-
Tail	<u>Med2/29</u>	Med2	Hintersex	-
	<u>Med3/27</u>	Med3/Pgd1/Hrs1	TRAP37/CRSP347	-
	Med5/24	Nut1	TRAP100/DRIP100/CRSP100	-
	<u>Med15</u>	Gal11	ARC105	2KO4/2K0N/2GUT
	Med16	Sin4	TRAP95/DRIP92	-
	Med23	-	TRAP150 $\beta$ /DRIP130/CRSP130/hSur2	-
Kinase	Med12	Srb8	TRAP230/DRIP240	-
	Med13	Srb9	TRAP240/DRIP250	-
	Cdk8	Srb10/Ssn3/Ume5	hSrb10/CDK8	-
	CycC	Srb11/Ssn8/Ume3	hSrb11/CycC	1ZP2
Unassigned	Med25	-	ARC92/ACID1	2XNF/2L6U/2L23
	Med26	-	ARC70/CRSP70	-
	Med28	-	Fksg20	-
	Med30	-	TRAP25	-

<sup>a</sup> Subunits comprising the ancient core Mediator are shown in red (Bourbon, 2008); subunits that are essential for yeast viability are shown are underlined.

<sup>b</sup> pdb code is given when structural information on individual subunits or subcomplexes is available

### **1.4.2 Modular architecture and structure**

Based on electron microscopy (Asturias et al, 1999; Davis et al, 2002; Dotson et al, 2000), biochemical studies (Kang et al, 2001) and gene expression profiling (van de Peppel et al, 2005), Mediator subunits were suggested to reside in four flexibly linked modules, the head, middle, tail and the dissociable kinase modules (Figure 3). The modular architecture as well as the subunit composition appears to be conserved from yeast to man (Table 3).

The Sc Mediator complex has a molecular weight of around 1.4 MDa and comprises 25 subunits, which have all homologues in higher eukaryotes (Bourbon, 2008). 10 subunits, which are all part of the core Mediator, are essential for cell viability (Table 3) (Myers & Kornberg, 2000). Many of the subunit-subunit interactions within Mediator are known from a large-scale yeast two-hybrid screen (Guglielmi et al, 2004) and biochemical studies (Baumli et al, 2005; Beve et al, 2005; Koschubs et al, 2010; Larivière et al, 2006; Takagi et al, 2006). Structural information on Mediator subunits is available for CycC (Hoepfner et al, 2005), Med7C/21 (Baumli et al, 2005), Med7N/31 (Koschubs et al, 2009), Med8C/18/20 (Larivière et al, 2006), Med15 (Thakur et al, 2008; Yang et al, 2006) and Med25 (Bontems et al, 2010) (Table 3). The architecture of the Mediator head (Takagi et al, 2006) and middle module (Baumli et al, 2005; Koschubs et al, 2010) is well established through recombinant coexpression and copurification studies. The two modules were suggested to interact through the head module subunit Med6 and the middle module subcomplex Med7C/21 (Baumli et al, 2005). Numerous low resolution electron microscopic reconstructions of the complete Mediator from various species revealed the highly flexible nature of the complex and large conformational changes in response to activator binding (Asturias et al, 1999; Cai et al, 2009; Cai et al, 2010; Davis et al, 2002; Dotson et al, 2000; Elmlund et al, 2006; Näär et al, 2002; Taatjes et al, 2002; Taatjes et al, 2004).

### **1.4.3 Mediator function in transcriptional regulation**

Mediator function is best described as integrating signals from various regulatory proteins and transferring a calibrated output to the basal transcription machinery (Bjorklund & Gustafsson, 2005; Kornberg, 2005; Malik & Roeder, 2010; Naar et al, 2001). Mediator interacts with many medically important human transcription regulators, including hormone receptors (Taatjes et al, 2004), vitamin D receptor (Taatjes et al, 2004) and p53 (Meyer et al, 2010). Recent studies have linked Mediator

to various human diseases, like Alzheimer disease (Xu et al, 2011), congenital malformations (Muncke et al, 2003), mental retardation (Philibert & Madan, 2007) and cancer (Firestein et al, 2008; Morris et al, 2008; Zhu et al, 1999). In yeast, Mediator is globally required for Pol II transcription (Holstege et al, 1998). Mediator is thought to act by promoting transcription initiation complex assembly through activator-Mediator, Mediator-Pol II and Mediator-GTF contacts (Cantin et al, 2003). Mediator is recruited to the promoter by activators independent of Pol II recruitment (Cosma et al, 2001) and remains in the scaffold complex after initiation to enable rapid reinitiation of previously transcribed genes (Yudkovsky et al, 2000). A recent study has demonstrated a direct interaction between Mediator and cohesin at promoters in embryonic stem cells (Kagey et al, 2010). This interaction promotes gene looping and thus physically connects distant enhancer elements with the core promoter. The dissociable Mediator kinase module, comprising the Cdk8 kinase, CycC, Med12 and Med13, phosphorylates the CTD at Ser5 together with Kin28 (TFIIH) and consequently facilitates transcription initiation and promoter escape (Liu et al, 2004).

Global gene expression studies of Mediator deletion mutants have implicated the different modules in the regulation of different subsets of genes (Beve et al, 2005; Holstege et al, 1998; Singh et al, 2006; van de Peppel et al, 2005). The middle module is required for regulating HSP genes and low-iron response genes, the tail module for regulating HSP and OXPHOS genes, and the kinase module for regulating genes required during nutrient starvation.

#### **1.4.4 Head module architecture and function**

The *Sc* head module comprises five essential subunits (Med6, Med8, Med11, Med17, and Med22) and two non-essential subunits (Med18 and Med20) (Table 3). The head module architecture was characterized in detail by yeast-two-hybrid screens (Esnault et al, 2008; Guglielmi et al, 2004) as well as coexpression and copurification studies (Baumli et al, 2005; Larivière et al, 2006; Takagi et al, 2006). Med6, Med8, Med11 and Med22 interact directly with the central scaffolding subunit Med17, while the dimeric Med18/20 subcomplex interacts with the C-terminus of Med8 (Med8C). Structural information is only available for this non-essential Med8C/18/20 subcomplex (Larivière et al, 2006).

The Mediator head module plays an important role during PIC assembly, contacting Pol II-TFIIF (Takagi et al, 2006), Pol II CTD (Kang et al, 2001; Näär et al, 2002), TATA-

binding protein (TBP) (Kang et al, 2001; Larivière et al, 2006), TFIIB (Kang et al, 2001) and TFIIH (Esnault et al, 2008). A temperature-sensitive mutation in Med17 (*med17-ts*; *srb4-138* strain) affecting head module integrity abolishes the stimulation of basal transcription *in vitro* (Ranish et al, 1999; Takagi et al, 2006), prevent association of Pol II and GTF with core promoters *in vivo* (Bhaumik et al, 2004) and cause a global shutdown of mRNA synthesis (Holstege et al, 1998; Thompson & Young, 1995). *Sc* Med11 and Med22 are important for head module architecture and function. They both bind and stabilize the central head subunit Med17 (Kang et al, 2001; Takagi et al, 2006). A temperature-sensitive mutation in Med22 (*srb6-107*) causes similarly to *med17-ts* a global decrease in mRNA synthesis (Thompson & Young, 1995). Single point mutations in Med22 can act as extragenic suppressors of the *med17-ts* phenotype (Lee et al, 1998), of the cold-sensitive phenotype of Rpb1 CTD truncation (Thompson et al, 1993), and of the lethal CTD Ser2 phosphorylation site substitution mutation (Yuryev & Corden, 1996). Med11 was shown to directly interact with the Rad3 subunit of TFIIH in yeast-two-hybrid assays and a point mutation in Med11 reduced promoter occupancy of TFIIH kinase module (TFIIK) and consequently CTD Ser5 phosphorylation *in vivo* (Esnault et al, 2008). Despite the important role of Mediator head during transcription initiation, the underlying molecular mechanisms as well as the functions of individual subcomplexes are still unclear.

## 1.5 Aims and scope of this work

The general coactivator complex Mediator constitutes the central interface between activators and Pol II, enabling regulated transcription of most if not all protein-coding genes (Kornberg, 2005; Malik & Roeder, 2010). Despite its fundamental role in gene regulation, its three-dimensional structure and biochemical function are still poorly understood. The low abundance of Mediator in cells and its complex architecture made X-ray structure determination and high-resolution electron microscopy using natively purified complexes so far impossible. Nevertheless, only by correlating observed molecular interactions with functional roles *in vitro* and *in vivo*, the Mediator mechanism can be elucidated on a molecular level.

This study focused on the 7-subunit Mediator head module, which plays a pivotal role during transcription initiation. Conditional mutations in the head module abolish association of Pol II and the GTF with promoters *in vivo* (Bhaumik et al, 2004) and consequently cause a global shutdown of Pol II transcription (Holstege et al, 1998; Thompson & Young, 1995). Furthermore, several direct interactions with PIC components, namely Pol II-TFIIF (Takagi et al, 2006), Pol II-CTD (Kang et al, 2001; Näär et al, 2002), TATA-binding protein (TBP) (Kang et al, 2001; Larivière et al, 2006), TFIIB (Kang et al, 2001) and TFIIH (Esnault et al, 2008), had been described. Despite the importance of Mediator head, high-resolution structural information was only available for the non-essential Med8C/18/20 subcomplex. Therefore, a structure-function-system correlation, combining X-ray crystallography, yeast genetics, biochemical assays, chromatin immunoprecipitation, genome-wide expression profiling and label-free mass spectrometry, was used in this study to elucidate the submodular architecture and molecular mechanisms underlying Mediator head module function. In addition, *in vitro* transcription and PIC assembly assays using native yeast promoter templates, nuclear extracts and the corresponding yeast activator were established and optimized. These assays allow functional studies of Pol II transcription initiation in an *in vivo* like environment with the advantage of *in vitro* manipulations, like depletion and addition of recombinant factors. During the first main project, the function of the non-essential Med8C/18/20 subcomplex was characterized in detail and correlated with structural data already available in the lab (Larivière et al, 2006). In the second main project, the essential Med11/22 subcomplex, which interacts directly with TFIIH (Esnault et al, 2008), was structurally and functionally characterized.

## 1.6 Additional contributions

Additional experimental results leading to co-author publications are listed below. A more detailed description of the respective contributions is given in Chapter 3.4.

The Mediator middle submodule Med7N/31 was functionally characterized (Chapter 3.4.2). *In vitro* transcription with yeast nuclear extracts revealed a Med7N/31 requirement for activation of Pol II transcription by acidic activators. Furthermore, a cooperation of Med31 and TFIIS during transcription initiation was established (Koschubs et al, 2009).

The functional interaction of the metazoan Mediator subunit Med25 with the archetypical acidic transcription factor VP16 was analyzed (Chapter 3.4.3). An *in vitro* quenching assay was established to identify surface residues important for the Med25-VP16 interaction. Additionally, the individual contributions of VP16 subdomains and the underlying cooperativity during transcription activation in yeast were characterized. The results contribute to the understanding of how activation domains have evolved to adapt to different unrelated target surfaces (Vojnic et al., in press).

The functional role of TFIIB during scanning and transcription start site selection was analyzed (Chapter 3.4.4). *In vitro* assays with modified promoter templates were established to test mutations on the protein as well as on the DNA side. We could demonstrate that the INR alone determines TSS selection and that the -8 position is important for TSS recognition. Furthermore, our results suggest that the INR is recognized with the help of the TFIIB B-reader element during Pol II scanning (Kostrewa et al, 2009).

Furthermore, *in vitro* assays using yeast nuclear extracts were established to study the function of Iwr1 (Czeko et al., in press) and the Prp19 complex (Chanarat et al., manuscript in revision) during transcription. All results and experimental details are described in detail in the PhD thesis of the respective first author and the respective publication.

## 2 Materials and Methods

### 2.1 Materials

#### 2.1.1 Strains

**Table 4: Yeast strains (Sc)**

ID	Strain	Genotype	Source
272	wild-type (BY4741)	MATa; <i>his3Δ1</i> ; <i>leu2Δ0</i> ; <i>met15Δ0</i> ; <i>ura3Δ0</i>	Euroscarf (Y00000)
273	wild-type (BY4742)	MATa; <i>his3Δ1</i> ; <i>leu2Δ0</i> ; <i>lys2Δ0</i> ; <i>ura3Δ0</i>	Euroscarf (Y10000)
274	wild-type (BY4743)	MATa/MATα; <i>his3Δ1/his3Δ1</i> ; <i>leu2Δ0/leu2Δ0</i> ; <i>met15Δ0/MET15</i> ; <i>lys2Δ0/LYS2</i> ; <i>ura3Δ0/ura3Δ0</i>	Euroscarf (Y20000)
275	MED11/ <i>med11Δ</i>	BY4743; <i>YMR112C::KANMX4/YMR112C</i>	Euroscarf (Y26552)
276	MED22/ <i>med22Δ</i>	BY4743; <i>YBR253W::KANMX4/YBR253W</i>	Euroscarf (Y23393)
277	MED8 shuffle	MATa; <i>his3Δ1</i> ; <i>leu2Δ0</i> ; <i>met15Δ0</i> ; <i>ura3Δ0</i> ; <i>lys2Δ0</i> ; <i>YBR193C::KANMX4</i> ; pRS316 [LEU2 MED8]	Straesser Lab
278	MED11 shuffle	MATa; <i>his3Δ1</i> ; <i>leu2Δ0</i> ; <i>met15Δ0</i> ; <i>ura3Δ0</i> ; <i>lys2Δ0</i> ; <i>YMR112C::KANMX4</i> ; pRS316 [LEU2 MED11]	this study (MSeY059)
279	MED18 shuffle	MATa; <i>his3Δ1</i> ; <i>leu2Δ0</i> ; <i>met15Δ0</i> ; <i>ura3Δ0</i> ; <i>lys2Δ0</i> ; <i>YMR112C::KANMX4</i> ; pRS316 [LEU2 MED18]	Straesser Lab
280	MED22 shuffle	BY4741; <i>YBR253W::KANMX4</i> ; pRS316 [LEU2 MED22]	this study (MSeY041)
281	MED8 shuffle TAP-MED17	MATa; <i>his3Δ1</i> ; <i>leu2Δ0</i> ; <i>met15Δ0</i> ; <i>ura3Δ0</i> ; <i>lys2Δ0</i> ; <i>YBR193C::KANMX4</i> ; <i>MED17::TAP</i> ; pRS316 [LEU2 MED8]	Straesser Lab
282	MED11 shuffle MED7-TAP	MATa; <i>his3Δ1</i> ; <i>leu2Δ0</i> ; <i>met15Δ0</i> ; <i>ura3Δ0</i> ; <i>lys2Δ0</i> ; <i>YMR112C::KANMX4</i> ; <i>MED7::TAP::HIS3</i> ; pRS316 [LEU2 MED11]	this study (MSeY71)
283	MED11 shuffle KIN28-TAP	MATa; <i>his3Δ1</i> ; <i>leu2Δ0</i> ; <i>met15Δ0</i> ; <i>ura3Δ0</i> ; <i>lys2Δ0</i> ; <i>YMR112C::KANMX4</i> ; <i>KIN28::TAP::HIS3</i> ; pRS316 [LEU2 MED11]	this study (MSeY81)
284	MED11 shuffle RPB3-TAP	MATa; <i>his3Δ1</i> ; <i>leu2Δ0</i> ; <i>met15Δ0</i> ; <i>ura3Δ0</i> ; <i>lys2Δ0</i> ; <i>YMR112C::KANMX4</i> ; <i>RPB3::TAP::HIS3</i> ; pRS316 [LEU2 MED11]	this study (MSeY75)
285	MED18 shuffle TAP-MED17	MATa; <i>his3Δ1</i> ; <i>leu2Δ0</i> ; <i>met15Δ0</i> ; <i>ura3Δ0</i> ; <i>lys2Δ0</i> ; <i>YBR193C::KANMX4</i> ; <i>MED17::TAP</i> ; pRS316 [LEU2 MED8]	Straesser Lab
286	MED22 shuffle MED7-TAP	BY4741; <i>YBR253W::KANMX4</i> ; <i>MED7::TAP::HIS3</i> ; pRS316 [LEU2 MED22]	this study (MSeY70)
287	MED22 shuffle KIN28-TAP	BY4741; <i>YBR253W::KANMX4</i> ; <i>KIN28::TAP::HIS3</i> ; pRS316 [LEU2 MED22]	this study (MSeY80)
288	MED22 shuffle RPB3-TAP	BY4741; <i>YBR253W::KANMX4</i> ; <i>RPB3::TAP::HIS3</i> ; pRS316 [LEU2 MED22]	this study (MSeY74)
289	<i>med17-ts</i>	MATa; <i>his3Δ200</i> ; <i>leu2-3 112</i> ; <i>ura3-52Δ2</i> ; <i>SRB4::HIS3</i> ; pRY2882 [LEU2 <i>srb4-138</i> ];	Hahn Lab (SHY205)
290	<i>TFIIB-ts</i>	MATα; <i>ade-</i> ; <i>leu2Δ0</i> ; <i>SUA7::HIS3</i> ; p68.30 [LEU2 <i>sua7(G41E)</i> ]	Hahn Lab (SHY245)

**Table 5: Yeast strains (*Sp*)**

ID	Strain	Genotype	Source
291	wildtype (972h-)	<i>h</i> <sup>-</sup> ; <i>leu1-32</i> ; <i>ura4-Δ18</i> ;	H. Feldmann
292	MED7-TAP MED11-3HA	<i>h</i> <sup>-</sup> ; <i>leu1-32</i> ; <i>ura4-Δ18</i> ; <i>MED7::kanMX4::TAP</i> ; <i>MED11::ClonNAT::3HA</i>	this study

**Table 6: *E. coli* strains**

Strain	Genotype	Source
XL1-Blue	<i>rec1A</i> ; <i>endA1</i> ; <i>gyrA96</i> ; <i>thi-1</i> ; <i>hsdR17</i> ; <i>supE44</i> ; <i>relA1</i> ; <i>lac[F'</i> <i>proAB lacIqZΔM15 Tn10(Tet<sup>r</sup>)]</i>	Stratagene
BL21-Codon Plus (DE3) RIL	<i>B</i> ; <i>F</i> <sup>-</sup> ; <i>ompT</i> ; <i>hsdS</i> ( <i>r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup></i> ); <i>dcm</i> <sup>+</sup> ; <i>Tet<sup>r</sup></i> ; <i>gal λ</i> (DE3); <i>endA</i> ; <i>Hte</i> <i>[argU, ileY, leuW, Cam<sup>r</sup>]</i>	Stratagene
Rosetta B834 (RIL)	<i>F</i> <sup>-</sup> ; <i>ompT</i> ; <i>hsdSB</i> ( <i>r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup></i> ); <i>dcm</i> <sup>+</sup> ; <i>metB</i>	Novagen

## 2.1.2 Plasmids

**Table 7: Plasmids used for recombinant expression of *Med11/22* in *E. coli***

ID	Insert	Vector	Restriction sites	Source
1366	<i>Sc Med22</i> (1-121) RBS - <i>Sc Med11</i> ([-16] <sup>*</sup> -115)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP019)
1367	<i>Sc Med22</i> (1-121) RBS - <i>Sc Med11</i> ([-16] <sup>*</sup> -79)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP007)
1368	<i>Sc Med22</i> (1-121) RBS - <i>Sc Med11</i> ([-16] <sup>*</sup> -89)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP008)
1369	<i>Sc Med22</i> (54-121) RBS - <i>Sc Med11</i> ([-16] <sup>*</sup> -115)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP009)
1370	<i>Sc Med22</i> (1-121) RBS - <i>Sc Med11</i> (1-115)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP006)
1371	<i>Sc Med22</i> (1-121) RBS - <i>Sc Med11</i> (1-89)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP023)
1372	<i>Sc Med22</i> (54-121) RBS - <i>Sc Med11</i> (1-115)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP022)
1373	<i>Sc Med22</i> (1-89) RBS - <i>Sc Med11</i> (1-115)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP029)
1374	<i>Sc Med22</i> (54-89) RBS - <i>Sc Med11</i> (1-115)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP027)
1375	<i>Sc Med22</i> (1-121) RBS - <i>Sc Med11</i> (5-89)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP034)
1376	<i>Sc Med22</i> (1-121) RBS - <i>Sc Med11</i> (43-89)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP035)
1377	<i>Sc Med22</i> (1-89) RBS - <i>Sc Med11</i> (1-89)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP036)
1378	<i>Sc Med22</i> (1-89) RBS - <i>Sc Med11</i> (5-89)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP037)
1379	<i>Sc Med22</i> (Δ32-45) RBS - <i>Sc Med11</i> (5-89)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP049)

table continued on next page

<b>ID</b>	<b>Insert</b>	<b>Vector</b>	<b>Restriction sites</b>	<b>Source</b>
1380	Sc Med22 ( $\Delta$ 29-53) RBS - Sc Med11 (5-89)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP050)
1381	Sc Med22 (1-117) RBS - Sc Med11 (5-110)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP134)
1382	Sc Med22 (1-121) RBS - Sc Med11 (1-105)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP141)
1383	Sc Med22 (1-108) RBS - Sc Med11 (1-115)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP142)
1384	Sc Med22 (1-117) RBS - Sc Med11 (1-115)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP143)
1385	Sc Med22 (1-121) RBS - Sc Med11 (1-110)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP144)
1386	6His-Thrombin- Sc Med22 (1-121) RBS - Sc Med11 (1-115)	pET28b	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP151)
1387	6His-Thrombin- Sc Med22 (1-89) RBS - Sc Med11 (1-115)	pET28b	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP043)
1388	6His-Thrombin- Sc Med22 (54-89) RBS - Sc Med11 (1-115)	pET28b	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP042)
1389	6His-Thrombin- Sc Med22 (1-89) RBS - Sc Med11 (5-89)	pET28b	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP044)
1390	6His-Thrombin- Sc Med22 (54-89) RBS - Sc Med11 (5-89)	pET28b	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP045)
1391	6His-Thrombin- Sc Med22 (54-121) RBS - Sc Med11 (1-115)	pET28b	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP152)
1392	6His-Thrombin- Sc Med22 (1-117) RBS - Sc Med11 (1-115)	pET28b	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP153)
1393	6His-Thrombin- Sc Med22 (1-108) RBS - Sc Med11 (1-115)	pET28b	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP154)
1394	6His-Thrombin- Sc Med22 (1-121) RBS - Sc Med11 (43-115)	pET28b	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP155)
1395	6His-Thrombin- Sc Med22 (1-121) RBS - Sc Med11 (1-105)	pET28b	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP156)
1396	6His-Thrombin- Sc Med22 (1-121) RBS - Sc Med11 (1-110)	pET28b	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP157)
1397	6His-Thrombin- Sc Med22 (1-121) RBS - Sc Med11 (1-89)	pET28b	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP158)
1398	6His-Thrombin- Sc Med22 (1-89) RBS - Sc Med11 (1-89)	pET28b	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP159)
1399	Sc Med22 (1-121) RBS - Sc Med11 (1-115; T31A)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP062)
1400	Sc Med22 (1-121) RBS - Sc Med11 (1-115; L73P)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP061)
1401	Sc Med22 (1-121; I25K) RBS - Sc Med11 (1-115)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP123)
1402	Sc Med22 (1-121; I66K) RBS - Sc Med11 (1-115)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP124)
1403	Sc Med22 (1-121; N59H) RBS - Sc Med11 (1-115)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP125)
1404	Sc Med22 (1-121; N59E/N60E) RBS - Sc Med11 (1-115; T31A)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP126)
1405	Sc Med22 (1-121) RBS - Sc Med11 (1-115; E17K)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP128)
1406	Sc Med22 (1-121; N86K) RBS - Sc Med11 (1-115)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP129)

table continued on next page

ID	Insert	Vector	Restriction sites	Source
1407	Sc Med22 (1-121; I25K) RBS - Sc Med11 (1-115; V52D)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP130)
1408	Sc Med22 (1-121; I25K) RBS - Sc Med11 (1-115; V52D/G92S)	pRSF-Duet	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP131)
1409	6His-Thrombin- Sc Med22 (1-121) RBS - Sc Med11 (1-115; E17K/L24K)	pET28b	<i>NdeI/EcoRI</i> <i>EcoRI/XhoI</i>	this study (MSeP160)
1410	Ag Med11-TEV-6His (1-113) RBS - Ag Med22 (1-122)	pET24b	<i>NheI/SalI</i> <i>SalI/SacI</i>	this study (MSeP025)
1411	Ag Med11 (1-113) RBS - Ag Med22 (1-122)	pET24b	<i>NheI/SalI</i> <i>SalI/SacI</i>	this study (MSeP033)
1412	Ag Med11 (1-88) RBS - Ag Med22 (1-122)	pET24b	<i>NheI/SalI</i> <i>SalI/SacI</i>	this study (MSeP031)
1413	Cg Med11-TEV-6His (1-125) RBS - Cg Med22 (1-114)	pET24b	<i>NheI/SalI</i> <i>SalI/SacI</i>	this study (MSeP026)
1414	Cg Med11 (1-125) RBS - Cg Med22 (1-114)	pET24b	<i>NheI/SalI</i> <i>SalI/SacI</i>	this study (MSeP030)
1415	Cg Med11 (1-93) RBS - Cg Med22 (1-114)	pET24b	<i>NheI/SalI</i> <i>SalI/SacI</i>	this study (MSeP032)
1416	Sp Med11 (1-112) RBS - Sp Med22 (1-136)	pET28b	<i>NdeI/SalI</i> <i>SalI/SacI</i>	this study (LL324)
1417	Sp Med11 (1-91) RBS - Sp Med22 (1-136)	pET28b	<i>NdeI/SalI</i> <i>SalI/SacI</i>	this study (LL325)
1418	Sp Med11 (1-112) RBS - Sp Med22 (1-89)	pET28b	<i>NdeI/SalI</i> <i>SalI/SacI</i>	this study (LL326)

**Table 8: Plasmids used for recombinant expression of other proteins in *E. coli***

ID	Insert	Vector	Restriction sites	Source
1419	Sc Med17 (377-687)	pET21b	<i>SalI/NotI</i>	this study (MSeP145)
529	Sc Med17 (377-687) -6His	pET21b	<i>SalI/NotI</i>	this study (LL95)
670	Sp Med17 (1-545)	pCDF-Duet	<i>NcoI/NotI</i>	this study (LL208)
1420	Sp Med17 (257-545)	pCDF-Duet	<i>NcoI/NotI</i>	this study (LL338)
286	Sc Med8 (190-223) -6His RBS - Sc Med20 RBS - Sc Med18	pET24b		Lariviere et al., (2006)
235	Sc Med20 RBS - Sc Med18 -6His	pET24b		Lariviere et al., (2006)
473	Sp Med8 (180-200) -6His RBS - Sp Med18	pET24b		this study (LL16)
460	Sp Med18 -6His	pET24b		this study (LL3)
1421	Sc Gcn4 -6His	pET21a		Hahn laboratory (pJF28)
1422	6His- Gal4-VP16	pET21b	<i>NdeI/NotI</i>	this study
640	Gal4-AH	pET21b	<i>NheI/BamHI</i>	this study (LL170)

**Table 9: Plasmids used as templates for yeast complementation experiments**

ID	Insert	Vector	Restriction sites	Source
1423	Sc MED11 +500/300 bp up-/downstream DNA	pRS316	<i>SacII/ApaI</i>	this study (MSeP064)
1424	Sc MED11 +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP063)
1425	<i>Sc med11</i> ( $\Delta 2-42$ ) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP103)
1426	<i>Sc med11</i> ( $\Delta 106-115$ ) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP104)
1427	<i>Sc med11</i> ( $\Delta 80-115$ ) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP105)
1428	<i>Sc med11</i> ( $\Delta 90-115$ ) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP106)
1429	<i>Sc med11</i> ( $\Delta 111-115$ ) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP150)
1430	<i>Sc med11</i> (V52D/G92S) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP137)
1431	<i>Sc med11</i> (E17K/L24K) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP108)
1432	<i>Sc med11</i> (E17K/L20F/L24K) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP107)
1433	<i>Sc med11</i> (V52D) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP067)
1434	<i>Sc med11</i> (L73P) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP068)
1435	<i>Sc med11</i> (G92S) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP069)
1436	<i>Sc med11</i> (Y56K) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP071)
1437	<i>Sc med11</i> (L10K) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP086)
1438	<i>Sc med11</i> (E17K) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP087)
1439	<i>Sc med11</i> (L24K) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP088)
1440	<i>Sc med11</i> (Q25E) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP089)
1441	<i>Sc med11</i> (T31K) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP090)
1442	<i>Sc med11</i> (R39E) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP091)
1443	<i>Sc med11</i> (R67E) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP092)
1444	<i>Sc med11</i> (T31A) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP094)
1445	<i>Sc med11</i> (F55A) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP095)
1446	Sc MED11-3HA +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP165)
1447	<i>Sc med11</i> ( $\Delta 106-115$ ) -3HA +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP166)
1448	<i>Sc med11</i> (E17K/L24K) -3HA +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP167)

table continued on next page

ID	Insert	Vector	Restriction sites	Source
1449	<i>Sc</i> MED22 +500/300 bp up-/downstream DNA	pRS316	<i>SacII/ApaI</i>	this study (MSeP039)
1450	<i>Sc</i> MED22 +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP038)
1451	<i>Sc med22</i> ( $\Delta$ 2-53) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP040)
1452	<i>Sc med22</i> ( $\Delta$ 32-45) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP052)
1453	<i>Sc med22</i> ( $\Delta$ 29-53) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP051)
1454	<i>Sc med22</i> ( $\Delta$ 90-121) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP120)
1455	<i>Sc med22</i> ( $\Delta$ 118-121) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP146)
1456	<i>Sc med22</i> ( $\Delta$ 109-121) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP147)
1457	<i>Sc med22</i> (M56K) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP070)
1458	<i>Sc med22</i> (I25K) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP072)
1459	<i>Sc med22</i> (M27I) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP073)
1460	<i>Sc med22</i> (M63K) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP074)
1461	<i>Sc med22</i> (I66K) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP075)
1462	<i>Sc med22</i> (Q70E) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP076)
1463	<i>Sc med22</i> (L73K) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP077)
1464	<i>Sc med22</i> (R77E) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP078)
1465	<i>Sc med22</i> (K80E) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP079)
1466	<i>Sc med22</i> (E81K) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP080)
1467	<i>Sc med22</i> (L84K) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP081)
1468	<i>Sc med22</i> (L85K) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP082)
1469	<i>Sc med22</i> (N86K) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP083)
1470	<i>Sc med22</i> (Q70E) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP084)
1471	<i>Sc med22</i> (K9E) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP085)
1472	<i>Sc med22</i> (K67E) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP096)
1473	<i>Sc med22</i> (K9E/M63K) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP098)
1474	<i>Sc med22</i> (N59E/N60E) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP109)
1475	<i>Sc med22</i> (N59H) +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP122)

table continued on next page

ID	Insert	Vector	Restriction sites	Source
1476	<i>Sc med22 (V58L/N59H)</i> +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP121)
1477	<i>Sc med22 (K9E/E38K/M63T)</i> +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP138)
1478	<i>Sc med22 (L73E/K80E)</i> +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP148)
1479	<i>Sc med22 (K80E/L84E)</i> +500/300 bp up-/downstream DNA	pRS315	<i>SacII/ApaI</i>	this study (MSeP149)
1480	<i>Sc MED8</i> +500/300 bp up-/downstream DNA	pRS316		Straesser Lab
1481	<i>Sc MED8</i> +500/300 bp up-/downstream DNA	pRS315		Straesser Lab
1482	<i>Sc med8 (Δ190-223)</i> +500/300 bp up-/downstream DNA	pRS315		Straesser Lab

**Table 10: Plasmids used for *in vitro* assays**

ID	Insert	Vector	Restriction sites	Source
1483	Gal4 binding site + modified HIS4 core promoter + HIS4 INR	pBluescript II KS+	<i>PstI/BamHI</i>	Hahn Lab (pSH515)
1484	Gal4 binding site + modified HIS4 core promoter + HIS INR (A[-8]C)	pBluescript II KS+	<i>PstI/BamHI</i>	this study (MSeP020)
1485	Gal4 binding site + modified HIS4 core promoter + HIS INR (A[-9,-8,-7]C)	pBluescript II KS+	<i>PstI/BamHI</i>	this study (MSeP021)
1486	Gal4 binding site + modified HIS4 core promoter + HIS INR (CA[-1,+1]GC)	pBluescript II KS+	<i>PstI/BamHI</i>	this study (MSeP021)
1487	Gal4 binding site + modified HIS4 core promoter + 2x SNR14 INR	pBluescript II KS+	<i>PstI/BamHI</i>	this study (MSeP048)
1488	Gal4 binding site + modified HIS4 core promoter + 2x SNR14 INR (A[-8]C)	pBluescript II KS+	<i>PstI/BamHI</i>	this study (MSeP046)
1489	Gal4 binding site + modified HIS4 core promoter + 2x SNR14 INR (A[-8,-7]C)	pBluescript II KS+	<i>PstI/BamHI</i>	this study (MSeP041)
1490	Gal4 binding site + modified HIS4 core promoter (TATAmut) + 2x SNR14 INR	pBluescript II KS+	<i>PstI/BamHI</i>	this study (MSeP047)
792	native HIS4 promoter ([-428] - [+24] respective to the A in the start codon)	pBluescript II KS+	<i>HindIII/BamHI</i>	this study (LL279)
1492	native HIS4 promoter ([-428] - [+24] respective to the A in the start codon) (TATAmut)	pBluescript II KS+	<i>HindIII/BamHI</i>	this study (MSe161)
1493	native HIS4 promoter ([-428] - [+233] respective to the A in the start codon)	pBluescript II KS+	<i>HindIII/BamHI</i>	this study (MSe293)

### 2.1.3 Oligonucleotides

A detailed list with all oligonucleotide used for molecular cloning is available through the Cramer laboratory.

**Table 11: Oligonucleotides used for *in vitro* experiments**

Oligonucleotide	Sequence
EMSA pSH515-Gal4site-5'FAM fw	5'- <b>FAM</b> -GGGGATCGATCCGGGTGACAGCCCTCCGAATTCGAGCTCG
EMSA pSH515-Gal4site rv	5'-CGAGCTCGAATTCGGAGGGCTGTCAACCGGATCGATCCCC
EMSA pSH515-[-80TSS]-5'FAM fw	5'- <b>FAM</b> -CTCGAGAACAGTAGTATGCTG
EMSA pSH515-[-10TSS] rv	5'-GTATGTACAACACACATCGG
EMSA pSH515-[+20TSS] rv	5'-CTATTACACAGCGCAGTTGTG
Primer extension pSH515-5'FAM	5'- <b>FAM</b> -TTCACCAGTGAGACGGGCAACAGCCAAGCTC
Primer extension pSH515-5'Cy5	5'- <b>Cy5</b> -TTCACCAGTGAGACGGGCAACAGCCAAGCTC
Immobilized template pSH515 fw	5'- <b>Biotin</b> -TAATGCAGCTGGCAGCAGG
Immobilized template pSH515 rev	5'-GCCGCTCTAGCTGCATTAATG
Immobilized template LL279 fw	5'- <b>Biotin</b> -TAATGCAGCTGGCAGCAGG
Immobilized template LL279 rev	5'-GCCGCTCTAGCTGCATTAATG
Immobilized template LL293 fw	5'- <b>Biotin</b> -GGGTCAAGCACAAACAGCCGTG
Immobilized template LL293 rev	5'-TCCGGATCCGAACAGAGAAGAAACTCCGTTG

**Table 12: Oligonucleotides used for quantitative real-time PCR**

Oligonucleotide	Sequence
ILV5 promoter fw	ACCCAGTATTTTCCCTTTCC
ILV5 promoter rv	TTGTCTATATGTTTTGTCTTGC
ADH1 promoter fw	TTTCCTTCCTTCATTACGCACA
ADH1 promoter rv	TCAAGTAACTGGAAGGAAGGCCGTA
GAL1 promoter fw	GGGTAATTAATCAGCGAAGC
GAL1 promoter rv	GGTTATGCAGCTTTTCCATT
control region fw	TGCGTACAAAAAGTGTCAAGAGATT
control region rv	ATGCGCAAGAAGGTGCCTAT

## 2.1.4 Antibodies

**Table 13: Antibodies used in this study**

Antibody	Stock solution	Host	Source
anti-Med15	1:2000	rabbit	Abnova (PAB1161)
anti-Gcn4	1:400	rabbit	Santa Cruz (sc-50443)
anti-Med2	1:1000	goat	Santa Cruz (sc-28058)
anti-Rpb1-CTD	1:2000	mouse	Santa Cruz (sc-56767)
anti-Rpb3	1:1000	mouse	Neoclone (WP012)
anti-Med17	1:10000	rabbit	Hahn laboratory
anti-TBP	1:400	rabbit	Santa Cruz (sc-33736)
anti-TFIIB	1:4000	rabbit	Abcam (ab63909)
anti-HA-tag	1:700	rat	Roche (11867423001)
anti-flag-tag	1:1000	mouse	Sigma (F1804)
anti-TBP	1:400	rabbit	Santa Cruz (sc-33736)
anti-Taf4	1:5000	rabbit	Abcam (ab63910)
anti-Rpb11	1:1000	mouse	Neoclone (W0017)
Peroxidase Anti-Peroxidase (PAP)	1:2000	rabbit	Sigma (P1292)
anti-mouse-HRP	1:3000	goat	Bio Rad (170-6516)
anti-goat-HRP	1:3000	donkey	Santa Cruz (sc-2020)
anti-rabbit-HRP	1:3000	goat	Santa Cruz (sc-2004)
anti-rat-HRP	1:3000	goat	Sigma (A9037)

## 2.1.5 Growth media and additives

**Table 14: Growth media**

Growth medium	Description	Species
LB	1% (w/v) tryptone; 0.5% (w/v) yeast extract; 0.5% (w/v) NaCl; (+2% (w/v) agar for selective media plates)	<i>E. coli</i>
Minimal Medium	7.5 mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 8.5 mM NaCl; 55 mM KH <sub>2</sub> PO <sub>4</sub> ; 100 mM K <sub>2</sub> HPO <sub>4</sub> ; 1 mM MgSO <sub>4</sub> ; 20 mM glucose, 1 µg/l trace elements (Cu <sup>2+</sup> , Mn <sup>2+</sup> , Zn <sup>2+</sup> , Mo <sub>4</sub> <sup>2-</sup> ), 10 mg/l thiamine; 10 mg/l biotine; 1 mg/l Ca <sup>2+</sup> ; 1 mg/l Fe <sup>2+</sup> ; 100 mg/l amino acids (A, C, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, W, Y)	<i>E. coli</i>
YPD	2% (w/v) peptone; 2% (w/v) glucose; 1% (w/v) yeast extract; (+2% (w/v) agar for selective media plates)	<i>Sc</i>
Synthetic complete (SC)	0.67% (w/v) yeast nitrogen base; 0.06% (w/v) complete synthetic mix of amino acids; drop out as required; 2% (w/v) glucose; (+2% (w/v) agar for selective media plates)	<i>Sc</i>
5-FOA plates	SC (-ura); 0.01% (w/v) uracil; 0.2% (w/v) 5-FOA; 2% (w/v) agar for selective media plates	<i>Sc</i>
Sporulation plates	0.1% (w/v) yeast extract; 1% (w/v) potassium acetate; 0.079% (w/v) CSM amino acid complete mix; 0.25% (w/v) glucose; pH 5.6-6.0; 2% (w/v) agar for selective media plates	<i>Sc</i>
Rich medium	2% (w/v) peptone; 3% (w/v) glucose; 1% (w/v) yeast extract; (+2% (w/v) agar for selective media plates)	<i>Sp</i>

**Table 15: Growth media additives**

Additive	Stock solution	Working concentration
Ampicillin	100 mg/ml ampicillin in H <sub>2</sub> O	100 µg/ml
Chloramphenicol	50 mg/ml chloramphenicol in 100% ethanol p.a.	50 µg/ml
Kanamycin	30 mg/ml kanamycin in H <sub>2</sub> O	30 µg/ml
Tetracycline	12.5 mg/ml tetracycline in 70% (v/v) ethanol p.a.	12.5 µg/ml
Geneticin (G418)	200 mg/ml in H <sub>2</sub> O	200 µg/ml
Nourseothricin (clonNAT)	200 mg/ml in H <sub>2</sub> O	100 µg/ml
IPTG	1 M in H <sub>2</sub> O	0.5 mM

### 2.1.6 Buffers and solutions

Standard buffers and solutions were prepared as described (Sambrook & Russell, 2001).

**Table 16: General buffers and solutions**

Buffer	Description	Application
TFB-1	30 mM K acetate; 50 mM MnCl <sub>2</sub> ; 100 mM RbCl; 10 mM CaCl <sub>2</sub> ; 15% (v/v) glycerol; (pH 5.8 at 25°C)	competent <i>E. coli</i>
TFB-2	10 mM MOPS (pH 7.0 at 25°C); 10 mM RbCl; 75 mM CaCl <sub>2</sub> ; 15% (v/v) glycerol	competent <i>E. coli</i>
TBE	8.9 mM Tris/HCl; 8.9 mM boric acid; 2 mM EDTA (pH 8.0, 25°C)	Agarose gel electrophoresis
4x stacking gel buffer	0.5 M Tris/HCl (pH 6.8 at 25°C); 0.4% (w/v) SDS	SDS-PAGE
4x resolving gel buffer	3 M Tris/HCl (pH 8.9 at 25°C); 0.4% (w/v) SDS	SDS-PAGE
Running buffer	25 mM Tris; 0.1% (w/v) SDS; 250 mM glycine	SDS-PAGE
5x SDS sample buffer	250 mM Tris/HCl (pH 7.0 at 25°C); 50% (v/v) glycerol; 0.5% (w/v) bromophenol blue; 7.5% (w/v) SDS; 500 mM DTT	SDS-PAGE
Coomassie gel staining solution	50% (v/v) ethanol; 7% (v/v) acetic acid; 0.125% (w/v) Coomassie Brilliant Blue R-250	SDS-PAGE
Destaining solution	5% (v/v) ethanol; 7.5% (v/v) acetic acid	SDS-PAGE
Transfer buffer	25 mM Tris; 192 mM glycine; 20% ethanol	Western blotting
10x PBS	1.37 M NaCl; 27 mM KCl; 20 mM KH <sub>2</sub> PO <sub>4</sub> ; 10 mM Na <sub>2</sub> HPO <sub>4</sub> ·2 H <sub>2</sub> O;	Western blotting
WB blocking buffer	2% (w/v) milk powder in 1x PBS	Western blotting
Swelling buffer	200 mM Tris/HCl (pH 8.5 at 25°C); 2% (w/v) SDS	Edman sequencing
TAP lysis buffer	50 mM Tris/HCl (pH 7.5 at 25°C); 100 mM NaCl; 1.5 mM MgCl <sub>2</sub> ; 0.15% NP40	
100x protease inhibitor cocktail	0.028 mg/ml Leupeptin; 0.137 mg/ml Pepstatin A; 0.017 mg/ml PMSF; 0.33 mg/ml benzamidin; in 100% ethanol p.a.	various
TE	10 mM Tris-HCl; 1 mM EDTA; (pH 8.0 at 25°C)	various

**Table 17: Recombinant protein purification buffers**

Buffer	Description	Application
Buffer A	20 mM Tris/HCl (pH 8.0 at 25°C); 150 mM NaCl; 5 mM DTT	Sc Med8C/18/20, Sc Med18/20, Sp Med8C/18 purification
Buffer B	20 mM Tris/HCl (pH 8.0 at 25°C); 2 mM DTT	
Buffer A	20 mM Tris/HCl (pH 8.0 at 25°C); 150 mM NaCl; 2 mM DTT	Med11/22 purification
Buffer B	20 mM Tris/HCl (pH 8.0 at 25°C); 50 mM NaCl; 2 mM DTT	
Buffer A	20 mM phosphate buffer (pH 6.5 at 25°C); 150 mM NaCl; 10 mM $\beta$ -mercaptoethanol;	Med25-ACID purification
Buffer B	20 mM phosphate buffer (pH 6.5 at 25°C); 4 mM DTT	
Buffer C	2.5 mM phosphate buffer (pH 6.5 at 25°C); 100 mM NaCl; 8 mM DTT	
Buffer A	50 mM Tris/HCl (pH 8.0 at 25°C); 150 mM NaCl; 10 mM $\beta$ -mercaptoethanol	Med7N/31 purification
Buffer B	50 mM Tris/HCl (pH 8.0 at 25°C); 100 mM NaCl; 10 mM $\beta$ -mercaptoethanol	
Buffer C	50 mM MES (pH 6.5 at 25°C); 150 mM NaCl; 10 mM $\beta$ -mercaptoethanol	
Buffer A	10 mM Tris/HCl; 500 mM NaCl; 10% (v/v) glycerol; 10 mM imidazole; 10 $\mu$ M ZnSO <sub>4</sub> ; 10 mM $\beta$ -mercaptoethanol; (pH 8.0 at 25°C)	Gal4-VP16 & Gal4-AH purification
Buffer B	10 mM Tris/HCl; 100 mM NaCl; 10% (v/v) glycerol; 10 mM imidazole; 10 $\mu$ M ZnSO <sub>4</sub> ; 10 mM $\beta$ -mercaptoethanol; (pH 8.0 at 25°C)	
Buffer C	20 mM Hepes; 10% (v/v) glycerol; 10 $\mu$ M Zn acetate; 1 mM DTT; (pH 7.5 at 25°C)	
Buffer D	20 mM Hepes; 150 mM K acetate; 10% (v/v) glycerol; 10 $\mu$ M Zn acetate; 1 mM DTT; (pH 7.5 at 20°C)	
Buffer A	20 mM Tris/HCl (pH 8.0 at 25°C), 500 mM NaCl; 5 mM DTT	Gcn4 purification
Buffer B	20 mM HEPES (pH 7.5 at 25°C), 10 % Glycerol, 1 mM EDTA, 2 mM DTT	
Buffer C	20 mM HEPES (pH 7.5 at 25°C), 150 mM K acetate, 10 % Glycerol, 1 mM EDTA, 2 mM DTT	

**Table 18: Buffers for *in vitro* biochemical assays**

Buffer	Description	Application
Buffer A	50 mM Tris (pH 7.5 at 25°C), 20 mM EDTA (pH 8.0 at 25°C), 30 mM DTT	Yeast nuclear extract preparation
YPD/S	2% (w/v) peptone; 2% (w/v) glucose; 1% (w/v) yeast extract; 1 M sorbitol;	
Buffer B	18% (w/v) polysucrose 400; 10 mM Tris (pH 7.5 at 25°C); 20 mM K acetate; 5 mM Mg acetate; 1 mM EDTA (pH 8.0 at 25°C); 0.5 mM spermidine; 0.15 mM spermine; 3 mM DTT	
Buffer C	100 mM Tris (pH 8.0 at 25°C); 50 mM K acetate; 10 mM MgSO <sub>4</sub> ; 20% (v/v) glycerol; 2 mM EDTA (pH 8.0 at 25°C); 3 mM DTT	
Buffer D	20 mM Hepes; 10 mM MgSO <sub>4</sub> ; 1 mM EGTA; 20% (v/v) glycerol; 3 mM DTT; (pH 7.6 at 4°C)	
5x transcription buffer	100 mM HEPES (pH 7.6 at 25°C); 500 mM K acetate; 5 mM EDTA (pH 8.0 at 20°C); 25 mM Mg acetate	<i>in vitro</i> assays
5x annealing buffer	25 mM Tris (pH 8.3 at 25°C); 375 mM KCl; 5 mM EDTA (pH 8.0 at 20°C)	
formamide sample buffer	80% formamide, 25 mM EDTA, 1.5% bromophenolblue	
mini sequencing gel	8% acrylamide (Rotiphorese®, Roth), 1x TBE, 7 M urea, 0.05% APS, 0.05% TEMED;	
beads blocking buffer	1x transcription buffer; 60 mg/ml casein; 5 mg/ml polyvinylpyrrolidone; 2.5 mM DTT	

## 2.2 Methods

### 2.2.1 Molecular cloning

#### **Preparation of chemically competent *E. coli***

200 ml LB medium were inoculated with an overnight culture of the desired *E. coli* strain (Table 6) to a starting OD<sub>600</sub> of 0.05. Cells were grown at 37°C to an OD<sub>600</sub> of 0.5, then incubated on ice for 10 minutes. All following steps were carried out at 4°C. Cells were pelleted by centrifugation (3000 g / 10 min / 4°C) and washed with 50 ml TFB-1. Cells were again pelleted by centrifugation (3000 g / 10 min / 4°C) and resuspended in 4 ml of TFB-2. Cell suspension was aliquoted into tubes on dry ice and subsequently stored at – 80°C.

#### **Polymerase chain reaction (PCR)**

PCR primers usually contained a 5' overhang (5'-GGGCCCGGG-3') followed by the desired restriction site and 20-25 nt complementary to the target sequence. Hexahistidine (6His) tags and thrombin cleavage sites were introduced either by in-frame cloning into the respective expression vector or by PCR. Bicistronic expression constructs were generated as described (Larivière et al, 2006). PCR reactions were carried out with Taq polymerase (Fermentas), Herculase or Herculase II polymerases (Stratagene), Pwo SuperYield DNA Polymerase (Roche) or Phusion High-Fidelity DNA Polymerase (Finnzymes) depending on the requirements and according to the manufacturer's manual. 50 µl PCR reactions typically contained 1-50 ng plasmid template or 100-500 ng yeast genomic DNA. Thermocycling programs usually comprised 30 cycles and were done on the Biometra T3000 Thermocycler. Annealing temperature and elongation times were adjusted to the specific needs of the individual reactions. PCR products were visualized by agarose gel electrophoresis (1-2% w/v agarose, 1:10000 SYBR Safe DNA gel stain (Invitrogen) and 1x TBE buffer).

#### **Mutagenesis**

Point mutations and deletions were introduced using the PCR overlap extension method (Higuchi et al, 1988). Forward and reverse mutagenesis primers were designed carrying the mutation flanked by 20-25 nt complementary sequence on both sides. First two standard PCR reactions were done, amplifying the 5' and 3' end of the gene with an overlapping region containing the desired mutation. In a second step the

two overlapping fragments were used as templates to amplify the whole gene carrying the newly introduced mutation.

### **Restriction digest, dephosphorylation and ligation**

PCR products were purified using the QIAquick PCR purification kit (Qiagen). Vectors and purified PCR product were typically digested over night using restriction endonucleases (New England Biolabs and Fermentas) according to the manufacturer's manual. Digested vectors were subsequently dephosphorylated by addition of 1 u FastAP enzyme (Fermentas) according to the manufacturer's manual. Typically, DNA fragments were separated by agarose gel electrophoresis (1-2% w/v agarose, 1:10000 SYBR Safe DNA gel stain (Invitrogen) and 1x TBE buffer) and extracted using the QIAquick Gel Extraction Kit (Qiagen). PCR products and linearized vectors were typically ligated overnight at 16°C using 5 u of T4 DNA ligase (Fermentas) in a 20 µl reaction. A 5- to 10-fold molar excess of insert relative to the linearized vector was used.

### **Transformation in *E. coli* and plasmid preparation**

Chemically competent *E. coli* XL-1 blue cells were transformed with DNA using a heat shock protocol. A 50 µl aliquot of cells was thawed on ice and either 5 µl of a ligation reaction or 1 µl of plasmid were added. After 15 min of incubation on ice, a heat shock was applied (45 sec / 42°C) followed by 2-5 min incubation on ice. 200 µl of LB medium were added and cells were incubated at for 1 h at 37°C shaking vigorously. Afterwards, cells were plated on LB-agar plates containing the corresponding antibiotics for selection of transformed cells and incubated over night at 37°C. A single colony was used to inoculate 5 ml LB-medium containing the corresponding antibiotics. The culture was grown over night at 37°C and plasmids were isolated using the QIAprep Spin MiniPrep kit (Qiagen). Newly generated plasmids were verified by restriction digest and DNA sequencing (Eurofins or GATC).

## **2.2.2 General protein methods**

### **Protein expression and selenomethionine labeling**

Proteins were recombinantly expressed in *E. coli* BL21-Codon Plus (DE3)-RIL cells transformed with the respective expression plasmids as described previously (see 2.2.1). LB medium, containing the corresponding antibiotics, was inoculated with a pre-

culture to a starting OD<sub>600</sub> of 0.1. Cultures were grown to an OD<sub>600</sub> of 0.5-0.8 (37°C / 140 rpm), then cooled down on ice for 30 min. Protein expression was induced by addition of 0.5 mM IPTG and incubation over night (18°C / 140 rpm). Cells were harvested by centrifugation (4400 g / 15 min / 4°C / Sorvall SLC6000 rotor). Cell pellets were flash-frozen in liquid N<sub>2</sub> and stored at -80°C.

Selenomethionine labeling of proteins was essentially done as described (Budisa et al, 1995). Proteins were recombinantly expressed in methionine auxotroph *E. coli* Rosetta B834 (DE3) cells. LB medium, containing the corresponding antibiotics, was inoculated with a pre-culture to a starting OD<sub>600</sub> of 0.1. Cultures were grown to an OD<sub>600</sub> of 0.6 (37°C / 140 rpm). Cells were then harvested (4400 g / 15 min / 4°C / Sorvall SLC6000 rotor), washed and resuspended in minimal medium containing 50 mg/l selenomethionine and the corresponding antibiotics. Cultures were grown to OD<sub>600</sub> 0.8 (37°C / 140 rpm) in the minimal medium. After cooling the culture down, protein expression was induced by addition of 0.5 mM IPTG and incubation over night (18°C / 140 rpm). Cells were harvested by centrifugation (4400 g / 15 min / 4°C / Sorvall SLC6000 rotor). Cell pellets were flash-frozen in liquid N<sub>2</sub> and stored at -80°C.

### **Cell lysis**

Cell pellets were thawed on ice and then resuspended in the appropriate lysis buffer containing protease inhibitor cocktail. Cells were disrupted on ice by sonication (12 min / 25% duty cycle / 30-40 output value). Afterwards the lysate was cleared by centrifugation (max. speed / 30 min / 4°C / Sorvall SS34 rotor).

### **Protein purification**

Typically, recombinant proteins were purified using an initial affinity column followed by an ion exchange column and a size exclusion column. The specific purification protocols vary for each protein/subcomplex (see Chapter 2.2.3.).

### **Protein concentration**

Purified proteins were concentrated using “Amicon Ultra” spin concentrators (Millipore) with appropriate sample volume and molecular weight cut-off. The manufacturer’s manual was followed.

**Protein concentration determination**

Protein concentration was usually determined by the Bradford assay (Bradford, 1976). The assay was performed according to the manufacturer's manual (Bio-Rad). A standard curve was generated for each batch of dye reagent using serum bovine albumin (Roth). Alternatively, protein concentrations were calculated from the absorbance at 280 nm measured with a NanoDrop spectrophotometer. Absorption coefficients were calculated using ProtParam (<http://expasy.org/tools/protparam.html>).

**SDS-polyacrylamide gel electrophoresis (SDS-PAGE)**

Protein samples were analyzed by SDS-PAGE (Laemmli, 1970) with 15-17 % acrylamide gels (acrylamide:bisacrylamide ratio = 37.5:1) in the Bio-Rad gel systems. Buffers and gel solutions are listed in Table 16. Before loading, SDS sample buffer was added and samples boiled for 2 min at 95°C. Samples requiring high resolution over a broad molecular weight range, like Mediator complex purifications and immobilized template assay elutions, were separated using pre-casted NuPAGE Novex Bis-Tris minigels (Invitrogen). MOPS and MES running buffers (Invitrogen) were used according to the manufacturer's manual.

Gels were stained with Coomassie gel staining solution for approximately 1 h and subsequently destained with destaining solution over night. Alternatively, gels were used for Western blotting as described below.

**Western Blot**

Proteins were first separated by SDS-PAGE as described above and then transferred to a PVDF membrane (Schleicher & Schuell; pre-wet with 100% ethanol) using the wet blotting system from Bio-Rad. The manufacturer's manual was followed. Transfer was done either at 100 V for 1 h on ice or at 35 V over night at 4°C. For pre-cast NuPAGE Novex Bis-Tris minigels (Invitrogen) the recommended NuPAGE transfer buffer containing 20% (v/v) ethanol was used. After transfer, the membranes were air-dried for 30 min and then blocked for at least 1 h with WB blocking buffer. The blot was then incubated for at least 3 h at room temperature with the primary antibody in WB blocking buffer followed by three washing steps with WB blocking buffer, incubating 10 min at each step. Afterwards, the membrane was incubated with the secondary antibody in WB blocking buffer for at least 1.5 h. The membrane was washed three times with 1x PBS, incubating 10 min at each step. Secondary antibodies were usually coupled to

horseradish peroxidase. Signals were detected using the chemiluminescence kit (Pierce) followed by exposure of the membrane to high-sensitivity films (Invitrogen) and subsequent developing using a X-omat M35 developing machine (Kodak). Alternatively the chemiluminescent signals were detected with the Mini-LAS300 System (Fujifilm Life Sciences) and quantified using the ImageQuant software suite (GE Healthcare).

### **Edman sequencing**

Proteins were separated by SDS-PAGE, stained and destained as described above. The band of interest was excised, dried in a speed-vac and subsequently rehydrated in 50  $\mu$ l swelling buffer. Afterwards 200  $\mu$ l of ddH<sub>2</sub>O and a small piece of PVDF membrane (Schleicher & Schuell), pre-wet with 100% methanol, were added. Once the solution turned blue, 20  $\mu$ l of methanol were added. The sample was incubated approximately 2 days at room temperature. Then the membrane piece was washed 5x with 10% (v/v) methanol by incubating for 1 min each time. Afterwards, the protein was N-terminally sequenced from the dry membrane using a PROCISE 491 sequencer (Applied Biosystems).

### **Mass spectrometry**

Proteins were separated by SDS-PAGE, stained and destained as described above. For protein identification by mass spectrometry, the band of interest was excised, transferred to a tube containing 100  $\mu$ l ddH<sub>2</sub>O and sent to the protein analysis core facility (Adolf-Butenandt-Institute, LMU). Analysis of immobilized template eluates by mass spectrometry was performed in collaboration with Dr. Ignasi Forné (Group of Prof. Axel Imhof, LMU, Munich) as described in Chapter 2.2.4.

### **Limited proteolysis**

Limited proteolysis was typically done using chymotrypsin or trypsin proteases (Sigma). Usually, 20  $\mu$ g of purified protein in the respective gel filtration buffer were incubated with 100 ng of protease. Time courses were done by incubating the reactions at 37°C for 30 sec, 1 min, 3 min, 10 min, 30 min, and 60 min immediately followed by addition of SDS sample buffer and boiling for 2 min at 95°C. Fragments were separated by SDS-PAGE and identified by Edman sequencing as described above.

**Static light scattering**

Static light scattering measurements were done with a triple detector TDA (Viscotek) connected to a Superose-6 gelfiltration column (GE Healthcare) preequilibrated with corresponding buffer. 200  $\mu$ l of protein (approximately 2 mg/ml) were injected on the column and the eluate analysed by the triple detector, recording refractive index, UV and viscosity. From these values the hydrodynamic radius and the molecular weight were calculated using the static light scattering software package (Viscotek).

**Crystallization and structure solution**

Initial crystallization screens were performed at the MPI crystallization facility (MPI of Biochemistry, Martinsried). Typically, six different 96-well crystallization screens (Hampton Index, Nextal Classics, Nextal - The Cations, Nextal - The Anions, MPI Magic I and MPI Magic II) with 200 nl drop size were set up at room temperature by a robot. Fine screens around promising conditions were done manually in 24-well hanging drop plates (Easy Xtal Tool, Qiagen). Usually pH, protein concentration, precipitant and additive concentration was varied. Crystallization drops were set up in different protein:buffer ratios (1:1, 1:2 and 2:1) for each condition.

Crystals of *Schizosaccharomyces pombe* (*Sp*) Med18 and of the *Sp* Med8C/18 heterodimer were grown at 20°C in hanging drops over reservoirs containing 100 mM Tris (pH 8.5), 2 M sodium acetate, and 2 M sodium formate. The Med18 structure was solved by the single-wavelength anomalous dispersion method using selenomethionine labeled protein crystals. The Med8C/Med18 structure was solved by molecular replacement using the Med18 structure as a model .

Crystals of Med11/22 were grown at 20°C in hanging drops over reservoirs containing 100 mM MES pH 6, 5.5% PEG 6000 and 100 mM MgCl<sub>2</sub>. Microseeding was performed to optimize crystals. Initial crystals were transferred into 100  $\mu$ l reservoir solution and vortexed vigorously. 0.2  $\mu$ l of the resulting microseeding solution was added to each 2  $\mu$ l drop to nucleate crystal growth. Optimized crystals were harvested by gradually adding glycerol to a final concentration of 34% (v/v) and were subsequently flash-frozen in liquid nitrogen. Diffraction data was collected at 100 K on a PILATUS 6M detector at the Swiss Light Source SLS, Villigen, Switzerland. Diffraction data was processed using XDS and XSCALE (Kabsch, 1993). The program SOLVE (Terwilliger & Berendzen, 1999) identified 36 selenium sites in the asymmetric unit that were used for phasing. Solvent flattening, non-crystallographic symmetry averaging and initial

model building was done with RESOLVE (Terwilliger & Berendzen, 1999) and ARP/warp (Perrakis et al, 2001). The resulting electron density map allowed for manual building of most of Med11 and Med22 using COOT (Emsley & Cowtan, 2004). The model was refined using conjugate gradient minimization in PHENIX (McCoy et al, 2007). The asymmetric unit contained twelve Med11/22 heterodimers that deviated only slightly at the C-termini of the proteins. The structures and diffraction data have been deposited in the Protein Data Bank. All structure figures depict chain A/B and were prepared using PyMol (DeLano, 2002) .

### **2.2.3 Recombinant protein purification protocols**

Purification protocols for all recombinant proteins used in this study are listed below. All buffers are listed in Table 17.

#### **Med8C/18/20 and Med18/20 (Sc)**

The cleared lysate from two liters of culture was gradually precipitated by slowly adding saturated  $(\text{NH}_4)_2\text{SO}_4$  solution to a final concentration of 30%. Afterwards, suspension was stirred on ice for 30 min and centrifuged (15000 rpm / 30 min / 4°C / Sorvall SS34 rotor). Supernatant was discarded and the pellet resuspended in 8 ml of buffer A. Buffer B was added to a final conductivity of 10-15 mS/cm. The proteins were further purified by anion exchange chromatography using a MonoQ 10/100 column (GE Healthcare). The column was equilibrated with buffer B containing 100 mM NaCl and the complex was eluted with a linear gradient of 20 CVs from 100 mM to 500 mM NaCl in buffer B. Subsequently, the sample was applied to a Superdex 200 size exclusion column (GE Healthcare) equilibrated with buffer A. The protein was concentrated to approximately 1 mg/ml, flash frozen in small aliquots in liquid  $\text{N}_2$  and stored at -80°C.

#### **Med8C/18 (Sp)**

The cleared lysate from three liters of culture was loaded onto a 2 ml Ni-NTA column (Qiagen) equilibrated with buffer A. The column was washed with 10 CV of buffer A containing 10 mM imidazole and with 10 CV of buffer A containing 20 mM imidazole. The complex was eluted with buffer A containing 300 mM imidazole. Buffer B was added to a final conductivity of 10-15 mS/cm. The proteins were further purified by anion exchange chromatography using a MonoQ 10/100 column (GE Healthcare). The

column was equilibrated with buffer B containing 100 mM NaCl and the complex was eluted with a linear gradient of 10 CVs from 100 mM to 1 M NaCl in buffer B. After concentration the sample was applied to a Superose-12 size exclusion column (GE Healthcare) equilibrated with buffer A. The protein was concentrated to approximately 40 mg/ml for crystallization.

#### **Med11/22 (Sc)**

The cleared lysate from three liters of culture was loaded onto a 2 ml Ni-NTA column (Qiagen) equilibrated with buffer A. The column was washed with 10 CV of buffer A containing 10 mM imidazole and with 10 CV of buffer A containing 20 mM imidazole. The complex was eluted with buffer A containing 300 mM imidazole followed by overnight cleavage with thrombin while dialyzing against buffer B. The proteins were further purified by anion exchange chromatography using a HiTrap Q HP column (GE Healthcare). The column was equilibrated with buffer B and the complex was eluted with a linear gradient of 25 CVs from 50 mM to 500 mM NaCl in buffer B. Subsequently, the sample was applied to a HiLoad Superdex-75 pg 26/60 size exclusion column (GE Healthcare) equilibrated with buffer A.

#### **Med25-ACID (Hs)**

The cleared lysate from four liters of culture was loaded twice onto a 2 ml Ni-NTA column (Qiagen) equilibrated with buffer A. The column was washed with 10 CV of buffer B containing 1 M NaCl, 10 CV of buffer A and 10 CV of buffer A containing 10 mM imidazole. The protein was eluted with 10 CV of buffer A containing 200 mM imidazole, subsequently diluted with 10 CV of buffer B containing 50 mM NaCl and further purified by cation exchange chromatography using a MonoS column (GE Healthcare). The column was equilibrated with buffer B and the complex was eluted with a linear gradient of 10 CVs from 0 mM to 1 M NaCl in buffer B. Subsequently, the sample was applied to a Superose 6 size exclusion column (GE Healthcare) equilibrated with buffer C. The protein was concentrated to approximately 5 mg/ml, flash frozen in small aliquots in liquid N<sub>2</sub> and stored at -80°C.

**Med7N/31 (Sc)**

The cleared lysate from 2 liters of culture was loaded onto a 3 ml Ni-NTA column (Qiagen) equilibrated with buffer A. The column was washed with 20 CV of buffer A containing 20 mM imidazole. The complex was eluted with buffer A containing 200 mM imidazole, followed by overnight cleavage with thrombin. The proteins were further purified by anion exchange chromatography using a MonoQ 10/100 column (GE Healthcare). The column was equilibrated with buffer B and the complex was eluted with a linear gradient of 20 CVs from 100 mM to 1 M NaCl in buffer B. After concentration, the sample was applied to a HiLoad Superdex-200 size exclusion column (GE Healthcare) equilibrated with buffer C.

**TFIIB (Sc)**

The cleared lysate from one liter of culture was incubated in batch (20 min / 4°C), with 2 ml Ni-NTA material (Qiagen) equilibrated with buffer A containing 5 mM imidazole, 0.2% Tween and protease inhibitor cocktail. After transferring the material to a gravity flow column, it was washed with 10 CV buffer A containing 10 mM imidazole. The protein was eluted with 3 CV of buffer A containing 200 mM imidazole. The sample was diluted with buffer B to a final conductivity of 50 mSi/cm and further purified by cation exchange chromatography using a MonoS 10/100 GL column (GE Healthcare). The column was equilibrated with buffer C and the complex was eluted with a linear gradient of 15 CVs from 100 mM to 1 M NaCl in buffer C. Subsequently, the sample was applied to a Superdex 75 10/300 GL size exclusion column (GE Healthcare) equilibrated with buffer D.

**Gal4-VP16**

The cleared lysate from three liters of culture was loaded twice onto a 2 ml Ni-NTA column (Qiagen) equilibrated with buffer A. The column was washed with 10 CV of buffer A, 10 CV of buffer B and 5 CV of buffer B containing 20 mM imidazole. The protein was eluted with 10 CV of buffer B containing 200 mM imidazole and further purified by anion exchange chromatography using a HiTrap Q HP column (GE Healthcare). The column was equilibrated with buffer C and the complex was eluted with a linear gradient of 10 CVs from 0 mM to 700 mM NaCl in buffer C. Subsequently, the sample was applied to a Superose 12 size exclusion column (GE Healthcare)

equilibrated with buffer D. The sample was concentrated to approximately 0.5 mg/ml, flash frozen in small aliquots in liquid N<sub>2</sub> and stored at -80°C.

#### **Gal4-AH**

The cleared lysate from three liters of culture was loaded twice onto a 2 ml Ni-NTA column (Qiagen) equilibrated with buffer A. The column was washed with 10 CV of buffer A, 10 CV of buffer B and 5 CV of buffer B containing 20 mM imidazole. The protein was eluted with 10 CV of buffer B containing 200 mM imidazole and further purified by anion exchange chromatography using a HiTrap SP column (GE Healthcare). The column was equilibrated with buffer C and the complex was eluted with a linear gradient of 10 CVs from 0 mM to 700 mM NaCl in buffer C. Subsequently, the sample was applied to a Superose 12 size exclusion column (GE Healthcare) equilibrated with buffer D. The sample was concentrated to approximately 0.5 mg/ml, flash frozen in small aliquots in liquid N<sub>2</sub> and stored at -80°C.

#### **Gcn4**

The cleared lysate from two liters of culture was loaded twice onto a 2 ml Ni-NTA column (Qiagen) equilibrated with buffer A. The column was washed three times with 10 CV of buffer A containing 20 mM imidazole. The protein was eluted with 10 CV of buffer A containing 500 mM imidazole, subsequently diluted 1:5 with buffer B and further purified by cation exchange chromatography using a HiTrap SP column (GE Healthcare). The column was equilibrated with buffer B and the complex was eluted with a linear gradient of 10 CVs from 0 mM to 1 M NaCl in buffer B. Subsequently, the sample was applied to a Superdex 200 size exclusion column (GE Healthcare) equilibrated with buffer C. The sample was concentrated to approximately 0.5 mg/ml, flash frozen in small aliquots in liquid N<sub>2</sub> and stored at -80°C.

### **2.2.4 *In vitro* biochemical assays**

Yeast nuclear extract preparation, *in vitro* transcription and immobilized template assay were done as described (Ranish & Hahn, 1991; Ranish et al, 1999) with some changes.

#### **Yeast nuclear extract preparation**

For yeast nuclear extract preparation three liters of *Sc* were grown to  $5 \times 10^7$  cells/ml in YPD containing 50 µg/ml ampicillin. Cells were harvested by centrifugation and the pellet was resuspended in 30 ml buffer A and incubated at 30°C for 15 min. Afterwards, cells were centrifuged and resuspended in 20 ml YPD/S, 3 ml 2 M sorbitol, 3 ml buffer A containing 18 mg zymolyase (Seikagaku) and protease inhibitor cocktail. Cells were incubated at 30°C for 15 – 60 min until at least 85% of cells were spheroblasted, then 100 ml of YPD/S was added and cells were centrifuged. Pellet was resuspended in 250 ml YPD/S and incubated at 30°C for 30 min. Afterwards cells were washed twice with 200 ml ice-cold YPD/S and once with 200 ml 1 M ice-cold sorbitol. Cells were resuspended in 100 ml ice-cold buffer B and lysed on ice in a Dounce glass homogenizer (Kontes) with a small pestle. Crude nuclei were isolated by centrifuging (5000 rpm / 4°C / SLA-1500 rotor) two times for 8 min and two times for 5 min, transferring the supernatant each time to a new bottle. Afterwards nuclei were pelleted by centrifugation (13000 rpm / 30 min / 4°C / SS34 rotor), washed once with 15 ml buffer C, resuspended in 15 ml buffer C, flash frozen in liquid N<sub>2</sub> and stored at -80°C. Nuclei were lysed by adding 3 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (pH 7.5 at 20°C) to a final concentration of 0.5 M and incubation (30 min / 4°C / turning wheel). Afterwards nuclear lysate was ultracentrifuged (28000 rpm / 90 min / 4°C). Nuclear proteins in the supernatant were precipitated by addition of 0.35 g solid (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> per ml and incubation (30 min / 4°C / turning wheel). After centrifugation the nuclear proteins were resuspended in 250-1000 µl of buffer D and dialysed for 4.5 h against buffer D containing 75 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>. Nuclear extracts were flash frozen in liquid nitrogen and stored at -80°C. Concentration of nuclear extracts was determined by Bradford assay diluting the sample appropriately with 0.1% (w/v) SDS.

**Pol II *in vitro* transcription with nuclear extracts**

Pol II *in vitro* transcription was done on various yeast promoter DNA templates inserted into pBluescript II KS+ plasmid (Table 9). The 25  $\mu$ l transcription reaction mixture contained 1x transcription buffer, 200  $\mu$ g yeast nuclear extract, 200 ng of template plasmid, 192  $\mu$ g of phosphocreatine, 0.2  $\mu$ g of creatine phosphokinase, 10 U of RiboLock RNase inhibitor (Fermentas) and 100  $\mu$ M nucleoside triphosphates (NTPs). For activated transcription, 150 ng of recombinant Gal4-VP16, 200 ng of Gal4-AH or 200 ng of recombinant Gcn4 was added. The reaction was incubated for 60 min at 18°C. For temperature-sensitive yeast nuclear extracts the reactions were incubated for 45 min at 30°C. RNA was isolated using the RNeasy MinElute kit (Qiagen) according to the manufacturer's manual. RNA was eluted from the column with 14  $\mu$ l RNase-free H<sub>2</sub>O and *in vitro* transcripts were subsequently analyzed by primer extension. The 20  $\mu$ l primer annealing reaction contained 1x annealing buffer, 12  $\mu$ l eluate from RNeasy MinElute column and 0.125 pmol fluorescently labeled oligo (Table 11). After boiling the samples 3 min at 95°C, the primer was annealed for 45 min at 48°C. Afterwards, 40  $\mu$ l synthesis mix containing 1x synthesis buffer, 0.15 mM dNTPs, 12.5 u MuLV reverse transcriptase (Roche) and 1  $\mu$ g actinomycin D, was added. The reverse transcription reaction was incubated for 30 min at 37°C. The resulting cDNA was EtOH precipitated and resuspended in 4  $\mu$ l RNase A (40  $\mu$ g/ml). After incubating 3 min at 18°C, 4  $\mu$ l formamide sample buffer were added and the samples were boiled for 1 min. Transcripts were separated on a mini sequencing gel in 1x TBE buffer, scanned with a Typhoon 9400 and quantified with the ImageQuant software (GE Healthcare).

**Pol II immobilized template assay with nuclear extracts**

Pol II immobilized template assays were done on various linear yeast promoter templates. Templates were amplified by PCR from the *in vitro* transcription template plasmids (Table 9) using a biotin labeled forward primer and a regular reverse primer (Table 11). Afterwards, the PCR products were purified with the QIAquick PCR purification kit (Qiagen) followed by phenol chloroform extraction. For each reaction approximately 4.5 pmol of biotin labeled template were coupled to 200  $\mu$ g of magnetic streptavidin beads (Dynabeads M-280, Invitrogen). Coupled beads were blocked for 15 min at 20°C with bead blocking buffer and subsequently for 15 min at 20°C with 1x transcription buffer containing 0.5 mM biotin. After washing with 1x transcription buffer,

beads were resuspended in 20  $\mu$ l 1x transcription buffer and 5 pmol recombinant Gcn4 were added. After incubating for 10 min at 20°C while shaking, the beads were washed twice with 1x transcription buffer and resuspended in 20  $\mu$ l 1x transcription buffer. A 100  $\mu$ l immobilized template reaction contained 20  $\mu$ l of prepared beads, 1 mg yeast nuclear extract, 5  $\mu$ g competitor DNA (HaeIII digested genomic *E. coli* DNA), 1x transcription buffer, 768  $\mu$ g of phosphocreatine, 0.8  $\mu$ g of creatine phosphokinase, 0.05% NP-40 and 2.5 mM DTT. Assembly was done for 1 h at 4°C while shaking. Templates were washed twice with 1x transcription buffer containing 2  $\mu$ g/ml competitor DNA, twice with 1x transcription buffer containing 500 mM K acetate and twice with 1x transcription buffer. Proteins were eluted by boiling beads in SDS-loading dye and subsequently analyzed by SDS-PAGE and Western blotting.

#### **Analysis of immobilized template assays by tandem mass spectrometry**

For tandem mass spectrometry analysis, 5 identical immobilized template reactions were prepared as described above and pooled after the last washing step. Proteins were eluted by boiling beads in SDS-loading dye and subsequently loaded on a 15% SDS-polyacrylamide gel. The proteins were run approximately 1 cm into the resolving gel. Then the gel was stained with Coomassie and destained as described above. Each lane was cut horizontally in eight equally sized gel slices. Afterwards, proteins were reduced, alkylated and digested with trypsin. Peptides were analyzed by tandem mass spectrometry using a LTQ Orbitrap (Thermo) in collaboration with the Prof. Axel Imhof laboratory (LMU, Munich). Proteins were identified using MASCOT ([www.matrixscience.com](http://www.matrixscience.com)) and quantified label-free using spectral count analysis (Zhu et al, 2010).

#### **Electrophoretic mobility shift assays**

Binding reactions (20  $\mu$ l) contained 5 pmol Gal4-VP16 and 5, 50 or 150 pmol ACID variants. We used 1 pmol of a DNA duplex containing a single Gal4-binding site (Table 11). After incubation for 20 min at room temperature, ACID variants were added accordingly, and the mixture was incubated for 10 min. Formed complexes were separated on 5% acrylamide gels in TGOE buffer (0.25 M Tris, pH 8.3, 1.9 M glycine). Bands were quantified with a Typhoon 9400 scanner and the ImageQuant Software (Amersham Biosciences).

## 2.2.5 Yeast methods

### Cryo-stocks of yeast strains

A single colony was restreaked on a YPD plate and incubated for 2 days at 30°C. All cells from the plate were transferred to 1 ml sterile 40 % v/v glycerol, vortexed and flash-frozen in liquid nitrogen. Cryo-stocks were stored at -80°C.

### Yeast transformation

50 ml of YPD medium were inoculated with a pre-culture grown over night to a starting OD<sub>600</sub> of 0.1 and grown to a final OD<sub>600</sub> of 0.8 (30°C / 160 rpm). Cells were harvested by centrifugation (2500 rpm / 5 min / 4°C) and washed with 25 ml ddH<sub>2</sub>O. Afterwards, the cells were washed with 1 ml 100 mM lithium acetate, then resuspended in 400 µl 100 mM lithium acetate and divided into 100 µl aliquots. After centrifugation (max. speed / 15 sec) the supernatant was discarded and 240 µl 50% (w/v) PEG 3350, 36 µl 1 M lithium acetate, 50 µl boiled salmon sperm DNA (2 mg/ml) and 34 µl DNA (1-5 µg linear DNA or 200 ng plasmid DNA) were added. The cells were vortexed vigorously for 1 min, incubated for 30 min at 30°C and then heat-shocked for 15 min at 42°C. Afterwards, cells were centrifuged (max. speed / 15 sec), the supernatant discarded and the pellet resuspended in 200 µl YPD medium. After recovery (30°C / 1 h / shaking), cells were centrifuged again and the pellet resuspended in 200 µl ddH<sub>2</sub>O. 50 µl of the suspension was plated on the respective drop-out plates and incubated for 2-3 days at 30°C.

### Sporulation and tetrad dissection

Freshly grown diploid cells were restreaked on sporulation plates and incubated for several days at 30°C. Sporulation was monitored under the light microscope. When at least 5% of the cells had sporulated, a loop full of cells was washed twice with 500 µl ddH<sub>2</sub>O and resuspended in 100 µl ddH<sub>2</sub>O. The outer cell wall of the tetrads was destroyed by adding 10 µl Glusulase (Perkin-Elmer). After 5 min incubation at room temperature, suspension was put on ice and 400 µl of ddH<sub>2</sub>O were added. Fraction of cell suspension was streaked out on a YPD plate and spores were dissected using a tetrad microscope. Tetrads with four growing spores were restreaked on YPD plates and corresponding drop out plates to check for marker segregation.

**Complementation assays**

To assess the phenotype of mutants, yeast complementation assays were performed. The respective shuffle strain, carrying a genomic knockout of the gene of interest and wild-type copy of the gene including promoter and terminator on plasmid pRS316, was used. Various constructs including the promoter and terminator region on plasmid pRS315, typically empty plasmid, wild-type gene and mutated gene, were transformed into the respective shuffle strain. After selection on drop-out plates, equal amounts of freshly grown yeast cells in SDC (-Ura/-Leu) were resuspended in water and ten-fold dilutions were spotted on 5-FOA and SDC (-Ura/-Leu) plates. Viable mutant strains were streaked twice on 5-FOA plates and then on SDC (-Leu). Equal amounts of freshly grown yeast cells in SDC (-Leu) were resuspended in water, ten-fold dilutions were spotted on YPD plates and plates were subsequently incubated at 30°C and 37°C to assess fitness and temperature sensitivity.

**Growth curves**

Growth curves were recorded by inoculating 50 ml YPD medium with a pre-culture grown over night to a starting OD<sub>600</sub> of 0.1. The culture was grown at 30°C / 160 rpm for at least 10 h. Growth was monitored by measuring the OD<sub>600</sub> of the culture every hour.

**Tandem affinity purification**

Tandem affinity purification (TAP) from 3 l of yeast culture ( $5 \times 10^7$  cells/ml) was essentially done as described (Puig et al, 2001).

**Denaturing protein extraction**

Denaturing protein extraction from yeast cells was carried out essentially as described (Knop et al, 1996). A 5 ml yeast culture was grown in YPD overnight (30°C / 160 rpm). 5 OD<sub>600</sub> of cells were washed with 1 ml ddH<sub>2</sub>O and subsequently resuspended in 500 µl ddH<sub>2</sub>O. After addition of 150 µl 1.85 M NaOH containing 7.5% (v/v) β-mercaptoethanol, the suspension was incubated for 20 min on ice. Proteins were precipitated by adding 150 µl of 55% trichloroacetic acid and incubation for another 20 min on ice. After centrifugation (max. speed / 20 min / 4°C / table top centrifuge), the supernatant was removed and 1 ml of acetone added. After centrifugation (max. speed / 10 min / 4°C / table top centrifuge) the pellet was air-dried. 100 µl of 2x SDS sample

buffer and 20  $\mu$ l of 1 M Tris base were added. Samples were boiled for 2 min at 95°C before loading on a SDS-PAGE.

### **Immunoprecipitation from whole cell extracts**

A 50 ml yeast culture was grown in YPD (30°C / 160 rpm) to an OD<sub>600</sub> of approximately 5. Cells corresponding to 160 OD<sub>600</sub> were harvested by centrifugation, washed once with 25 ml ddH<sub>2</sub>O and resuspended in 1 ml TAP lysis buffer containing protease inhibitor cocktail and 0.5 mM DTT. Cell suspension was transferred to a 2 ml tube and 1 ml of 0.5 mm Zirconia beads (Roth) were added. Cells were lysed in a mixer mill (Retsch) for 30 min at 4°C. Lysate was transferred to a new tube and centrifuged (max. speed / 30 min / 4°C). 100  $\mu$ l of the cleared lysate were set aside as input control. To the remaining sample the detergent Igepal CA-630 (Sigma) was added to a final concentration of 0.05%. 20  $\mu$ l of IgG-coupled magnetic beads (Invitrogen) were washed twice with TAP lysis buffer containing 0.05% Igepal CA-630. Washed beads were added to the lysate and incubated on a turning wheel for 45 min at 4°C. Magnetic beads were subsequently washed 5 times with 1 ml TAP lysis buffer containing 0.05% Igepal CA-630, protease inhibitor cocktail and 0.5 mM DTT. After removing the last wash, 30  $\mu$ l of 2x SDS sample buffer were added to the beads and samples were boiled for 2 min at 95°C.

### **Chromatin immunoprecipitation**

All chromatin immunoprecipitation experiments were performed in biological duplicates in YPD as previously described in detail (Aparicio et al, 2005; Mayer et al, 2010). Since the Med11/22 mutant strains displayed severe growth defects and larger cell sizes compared to wild-type, cells were counted instead of measuring optical density. To minimize the risk of acquiring a rescue mutation and ensure biological significance of our observations, the used biological duplicates were already separated before shuffling out the respective rescue plasmid and several rounds of selection. Phenotype and growth was monitored closely at each step. Overnight cultures were diluted in fresh medium to 1x10<sup>6</sup> cells/ml (40 ml cultures, 160 rpm shaking incubator, 30°C) and grown to mid-log phase (1x10<sup>7</sup> cells/ml) before formaldehyde crosslinking. Input and immunoprecipitated samples were assayed by quantitative real-time PCR to assess occupancy of proteins at three different promoters. Primer pairs directed against the promoter of the highly transcribed ILV5 gene, the housekeeping gene ADH1, the

glucose-repressed GAL1 gene as well as against a heterochromatic control region of chromosome V were designed (Table 12) and the corresponding PCR efficiencies determined. All primer pairs used in this study had PCR efficiencies in the range of 95-100%. 25  $\mu$ l PCR reactions contained 1  $\mu$ l DNA template, 2  $\mu$ l of 10  $\mu$ M primer pairs and 12.5  $\mu$ l iTaq SYBR Green Supermix (Bio-Rad). Quantitative PCR was performed on a CFX96 Real-Time System (Bio-Rad) using a 3 min denaturing step at 95°C, followed by 49 cycles of 30 s at 95°C, 30 s at 61°C and 15 s at 72°C. Threshold cycle (Ct) values were determined using the Ct determination mode "Regression" of Bio-Rad CFX Manager software package (Version 1.1). Fold enrichment over heterochromatic control region was determined and calculated as described (Aparicio et al, 2005).

### 2.2.6 Gene-expression profiling

#### Experimental setup for „Structure-system correlation defines the gene regulatory Mediator submodule Med8C/18/20“ (see Chapter 3.1)

Gene expression profiling was done in collaboration with the laboratory of Prof. F. Holstege (University Medical Center Utrecht). All strains except *med8C* $\Delta$  are as described (van de Peppel et al, 2005). *med8C* $\Delta$  is isogenic to S288c. Truncation was made by using the ADH1 terminator from pFA6a-13myc-kanMX6 (Longtine et al, 1998). All experiments were performed in SC medium with 2% glucose. For microarray analysis, two independent colonies were inoculated and overnight cultures were diluted in fresh medium to an OD<sub>600</sub> of 0.15 (60 ml cultures, 250 rpm shaking incubator 30°C). Cells were harvested by centrifugation (4000 rpm, 3 min) at an OD<sub>600</sub> of 0.6, and pellets were frozen in liquid nitrogen. The RNA reference was obtained as described (van de Peppel et al, 2003). Total RNA was prepared by hot phenol extraction. Additionally, RNA was treated with DNase (Qiagen) and cleaned up using the RNeasy kit (Qiagen). The mRNA was amplified by *in vitro* transcription using T7 RNA polymerase on 1  $\mu$ g of total RNA. During *in vitro* transcription, 5-(3-aminoallyl)-UTP (Ambion) was incorporated into the single-stranded cRNA. Cy3 or Cy5 fluorophores (Amersham) were coupled to 3  $\mu$ g of cRNA. Before hybridization, free dyes were removed using RNAClean (Agencourt), and the efficiency of cDNA synthesis and dye incorporation was measured using a spectrophotometer (SpectraMax190, Molecular Devices). C6-amino-linked oligonucleotides (70 nucleotides in length), the Yeast Genome ArrayReady (Operon) were purchased from Qiagen and were printed on Codelink slides following manufacturers instructions (GE Healthcare) with a MicroGrid II

(Apogent Discoveries) using 48-quill pins (Microspot2500; Apogent Discoveries). Each gene is represented twice and the arrays additionally contained 2838 control features for external control normalization and QC (van de Peppel et al, 2003). From each sample, 2  $\mu$ g cRNA (with a specific activity of 2–6% dye-labeled nucleosides) was hybridized, together with 2  $\mu$ g cRNA from the reference, for 16–20 h at 42 °C (Agilent microarray hybridisation chamber). After scanning (G2565AA Agilent scanner, 100% laser power, 30% photomultiplier tube), raw data was extracted with Imagen 7.5 (Biodiscovery). After image quantification and local background subtraction, all negative values were replaced with the standard deviation of the local background. Print-tip with a span of 0.4 was applied to normalize genes (Yang et al, 2002). After averaging of duplicate spots for each gene on the array, data were visualized and clustered with GeneSpring 7.2 (SiliconGenetics). For each mutant individually, the replicate profiles were compared to the replicate wt profiles through the common reference with ANOVA (R/MAANOVA version 0.98-7 <http://cran.r-project.org/src/contrib/Descriptions/maanova.html>). In a fixed effect analysis, sample, array, spot, and dye effects were modeled. P-values were determined by a permutation F2 test in which residuals were shuffled 5000 times, Benjamini-Hochberg multiple-testing correction was applied. Genes with  $p < 0.05$  and an average fold change over the four measurements of at least 1.7 were considered significant. Pearson's correlation was calculated in Microsoft Excel. For each pair of deletion strains the respective lists of significantly changed genes were merged and the respective correlation coefficient (R-value) was calculated.

Overrepresented biological processes for genes with significant expression changes were determined using GO Slim Mapper (<http://db.yeastgenome.org/cgi-bin/GO/goSlimMapper.pl>) based on the GO database (Ashburner et al, 2000). GO Slim Mapper distributes genes to 32 different biological processes according to their GO annotation. Additionally it provides the corresponding genomic background level for each process. All genes categorized with the GO term “Biological process” or “not mapped” were listed together as “not annotated”.

**Experimental setup for „Mediator head subcomplex Med11/22 contains a common helix bundle building block with a specific function in transcription initiation complex stabilization“ (see Chapter 3.2)**

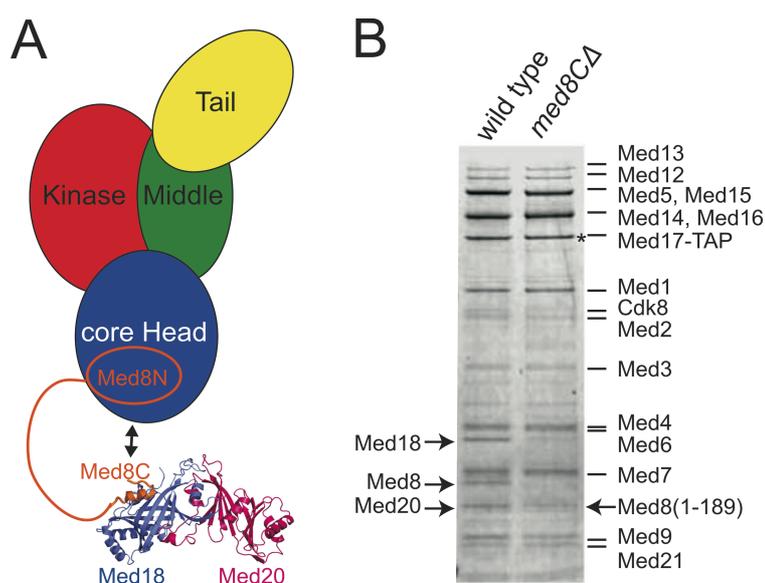
Gene expression profiling was done in the laboratory of Prof. P. Cramer. All experiments were performed in YPD medium with 2% glucose. For microarray analysis, at least two independent colonies were used for inoculation and overnight cultures were diluted in fresh medium to  $1 \times 10^6$  cells/ml (25 ml cultures, 160 rpm shaking incubator, 30°C). Cells were harvested in early log-phase ( $1 \times 10^7$  cells/ml) by centrifugation. Total RNA was prepared after cell lysis using a mixer mill (Retsch) and subsequent purification using the RNAeasy kit (Qiagen). The total RNA preparation was treated on-column with DNase (Qiagen). All following steps were conducted according to the Affymetrix GeneChip Expression Analysis Technical Manual (P/N 702232 Rev. 2). Briefly, one-cycle cDNA synthesis was performed with 1 µg of total RNA. *In vitro* transcription labeling was carried out for 16 h. The fragmented samples were hybridized for 16 h on Yeast Genome 2.0 expression arrays (Affymetrix), washed and stained using a Fluidics 450 station, and scanned on an Affymetrix GeneArray scanner 3000 7G. Data analysis was performed using R/Bioconductor (Gentleman et al, 2004). *S. pombe* probes were filtered out prior to normalization with the GCRMA algorithm (Wu et al, 2004). Linear model fitting and multiple testing correction using an empirical Bayes approach was performed using the LIMMA package (Smyth, 2004). Differentially expressed genes were defined as having an adjusted p-value smaller than 0.05 and an estimated fold change of at least 2.0 (calculated as the fold change of the average expression in the replicate measurements). Hierarchical clustering was calculated using TIGR MeV application (Saeed et al, 2003). Microarray data were submitted to the ArrayExpress database (<http://www.ebi.ac.uk/microarray>). Pearson's correlation coefficients were calculated as described above. Overrepresented biological processes for genes with significant expression changes were determined using GO Term Finder tool (<http://go.princeton.edu/cgi-bin/GOTermFinder>) based on the GO database (Ashburner et al, 2000).

### 3 Results and Discussion

#### 3.1 Structure-system correlation defines the gene regulatory Mediator submodule Med8C/18/20

##### 3.1.1 Med8C/18/20 is a subcomplex of the Mediator head

Previous results had revealed that *Sc Med8* contains an essential N-terminal domain (Med8N, residues 1–137), followed by a nonessential linker (residues 138–189) and a C-terminal region that includes a  $\alpha$ -helix (Med8C, residues 190–223) (Larivière et al, 2006). Med8C had been proposed to tether the Med18/20 heterodimer to the essential part of the Mediator head (Figure 4A).



**Figure 4: Med8C/18/20 is a subcomplex of Mediator head.**

**(A)** Overview of the Mediator architecture with the four modules head, middle, tail, and kinase. The head module is separated in the core head, constituted of Med6, Med8N, Med11, Med17, Med22, and the nonessential subcomplex of Med8C, Med18, and Med20 (in orange, blue, and magenta, respectively). **(B)** Deletion of Med8C *in vivo* leads to loss of Med18/20 after Mediator purification. N-terminally TAP-tagged Med17 was purified from wild-type yeast or from cells expressing Med8C $\Delta$ . The EGTA eluate after purification was separated using 4%–12% discontinuous SDS-PAGE and was analyzed by mass spectrometry after Coomassie staining. The copurifying proteins were identified by mass spectrometry. Arrows mark Med8, Med8C $\Delta$ , Med18, and Med20. An asterisk indicates TAP-tagged Med17.

To test this, we asked whether Med8C tethers the Med18/20 heterodimer to Mediator *in vivo*. We isolated Mediator by tandem affinity purification (TAP) from yeast strains expressing a TAP-tagged head subunit, Med17, and identified the copurifying Mediator subunits by mass spectrometry (Figure 4B). The same purification from a strain expressing a truncated version of Med8 that lacked Med8C (*med8CΔ*) resulted in a very similar pattern of protein bands, except that Med18 and Med20 were missing (Figure 4B). Therefore, retention of Med18 and Med20 in the Mediator complex requires Med8C. The truncated Med8 variant was, however, present in the preparation, showing that Med8C is not required to retain Med8 in the Mediator. These data are consistent with interaction data derived by coexpression and two-hybrid analysis (Guglielmi et al, 2004; Takagi et al, 2006) and suggest that the head contains two structural subcomplexes, the core head, consisting of all essential head subunits or subunit domains (Med6, Med8N, Med11, Med17, and Med22) and the nonessential Med8C/18/20 subcomplex (Figure 4A).

**Table 19: Data collection and refinement statistics for the *Sp* Med8C/18 structure**

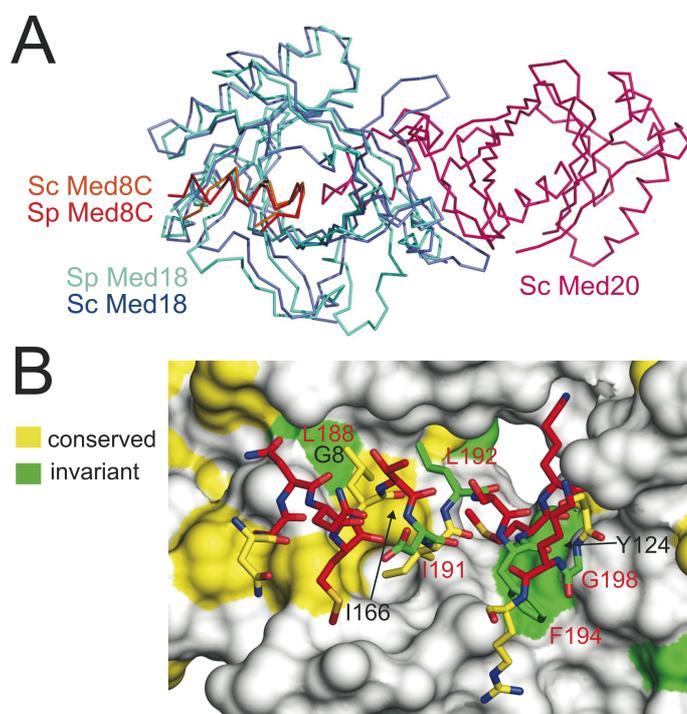
Crystal	SeMet-Med18	Med18	Med8C/18
Data collection			
Space group	C 2 2 2 <sub>1</sub>	C 2 2 2 <sub>1</sub>	P 3 <sub>2</sub> 2 1
Cell parameters			
a (Å)	81.8	81.8	111.0
b (Å)	97.1	97.2	111.0
c (Å)	129.8	130.0	68.6
γ (°)	90	90	120
Wavelength (Å)			
Wavelength (Å)	0.979	0.990	0.979
Resolution range (Å) <sup>a</sup>	20-2.9 (3.0-2.9)	20-2.7 (2.80-2.60)	20-2.4 (2.53-2.4)
Completeness (%)	99.6 (97.0)	99.1 (94.4)	99.4 (99.4)
Unique reflections	20,243 (1,967)	14,428 (1,344)	19,282 (2,806)
Redundancy	6.9 (4.9)	5.9 (4.7)	11.0 (11.2)
R <sub>sym</sub> (%)	8.5 (23.4)	5.6 (31.8)	8.4 (37.9)
<I>/<S>	18.6 (4.6)	28.3 (3.1)	21.4 (4.8)
Refinement			
Number of residues	224		
Number of non-hydrogen atoms	1,885		
Number of solvent molecules	62		
RMS bond deviation (Å)	0.006		
RMS angle deviation (°)	1.293		
Ramachandran plot (core/allowed/additionally allowed)	89.4/10.1/0.5		
R <sub>cryst</sub> (%)	21.1		
R <sub>free</sub> (%) <sup>b</sup>	23.1		

<sup>a</sup> The numbers in parenthesis correspond to the highest resolution shell

<sup>b</sup> 5% of the data were excluded from the refinement for free R-factor calculation

### 3.1.2 Conserved Med8C/18/20-core head interface

Because of the low sequence homologies between Mediator subunits from different species, the architecture of the Sc head could be a species-specific feature. To investigate this, we solved the crystal structure of the Med8C/18 complex from *Schizosaccharomyces pombe* (*Sp*). We coexpressed *Sp* Med18 with a hexahistidine-tagged *Sp* Med8C fragment corresponding to the *Sc* Med8C fragment used previously (Larivière et al, 2006) from a bicistronic vector in *E. coli*. The *Sp* Med8C fragment was sufficient for interaction with *Sp* Med18. The resulting stoichiometric Med8C/18 complex was crystallized and the structure solved (Table 19). *Sp* Med18 adopts a fold similar to its *Sc* ortholog (Figure 5A) with a root mean square deviation of 1.7 Å over 173 C $\alpha$  atoms.



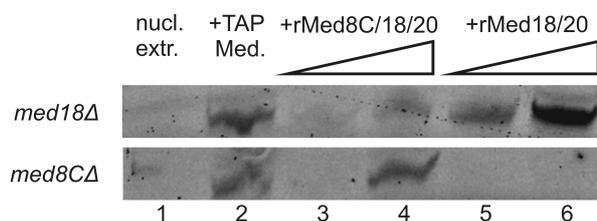
**Figure 5: Structural conservation of the Med8C/18 interaction**

**(A)** Superimposition of C $\alpha$  traces of Med8C/18 from *Sp* (in cyan and red; this study) and from *Sc* (in blue and orange; (Larivière et al, 2006)). Med20 is shown in magenta. **(B)** Interaction of *Sp* Med8C with *Sp* Med18. The Med8C helix is shown in red sticks, with conserved or invariant residues in contact with Med18 labeled in red. Med18 is shown as a white surface, with conserved or invariant residues in contact with Med8C labeled in black. Residues that are invariant and conserved between *Sc* and *Sp* are in green and yellow, respectively.

*Sp* Med8C forms a  $\alpha$ -helix, followed by a glycine-containing turn, and binds *Sp* Med18 across its central  $\beta$ -barrel as observed for its *Sc* counterpart (Figure 5A). Key contact residues in the Med8C–Med18 interface are conserved between *Sc* and *Sp* (Figure 5B). Given the large phylogenetic distance between these two fungi, the Med8C/18 interface is apparently also conserved in Mediator complexes of higher eukaryotes. Indeed, modeling of the human Med8C–Med18 interface showed that key contacts are conserved. Thus, the structural tethering of the Med18/20 heterodimer to the core head module through Med8C is conserved among eukaryotes.

### 3.1.3 Med8C/18/20 is required for activated transcription *in vitro*

To investigate whether the structural subcomplex Med8C/18/20 is also a functional subcomplex of Mediator, we conducted *in vitro* transcription assays. We prepared nuclear extracts from yeast strains carrying a deletion of the gene for Med18 (*med18* $\Delta$ ) or lacking the part of the Med8 gene coding for Med8C (*med8C* $\Delta$ ).



**Figure 6: Med8C/18/20 is required for activated transcription *in vitro*.**

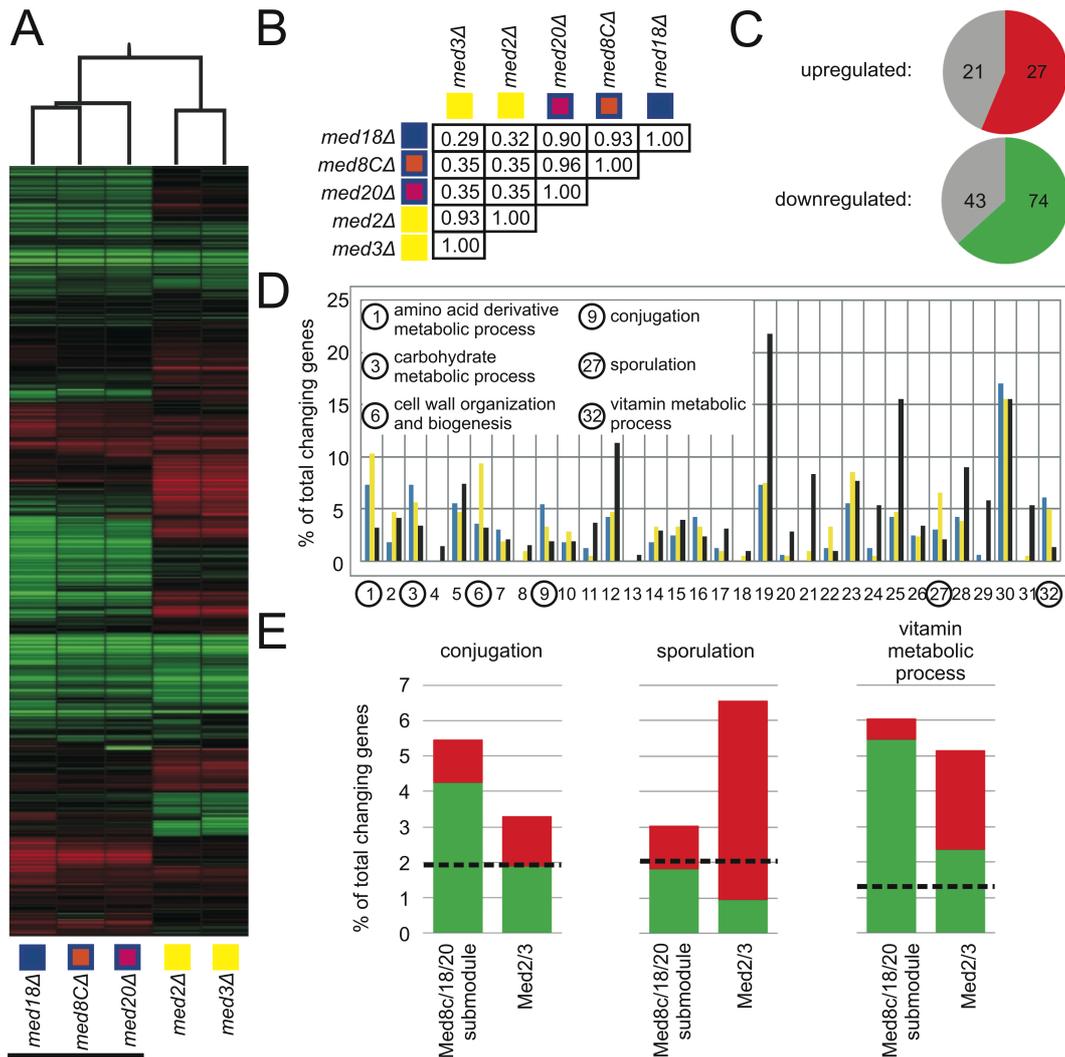
Med8C/18/20 is required for activated transcription *in vitro*. Assays were performed with *med18* $\Delta$  or *med8C* $\Delta$  nuclear extracts (lane 1), or extract to which 0.5 pmol TAP-purified Mediator was added, providing a positive control (lane 2), or extracts to which recombinant Med8C/18/20 (10x and 100x molar excess in lanes 3 and 4, respectively) or recombinant Med18/20 (10x and 100x molar excess in lanes 5 and 6, respectively) were added.

Consistent with previous data (Ranish et al, 1999; Thompson et al, 1993), a *med18* $\Delta$  nuclear extract did not support activated transcription (Figure 6, lane 1), apparently as a result of Mediator lacking Med18 and Med20 in this mutant. The transcription defect could indeed be rescued by addition of recombinant Med18/20 (Figure 6, lanes 5 and 6). This is consistent with the model that Med8C, which is present in the extract, tethers Med18/20 to the Mediator. Recombinant Med8C/18/20 subcomplex was far less efficient in rescue (Figure 6, lanes 3 and 4), likely because endogenous Med8C fails to replace recombinant Med8C for tethering Med18/20. Consistently, a nuclear extract

from the *med8C* $\Delta$  strain was inactive (Figure 6, lane 1), apparently as a result of Mediator lacking Med18 and Med20. Even a large excess of Med18/20 could not rescue the defect (Figure 6, lanes 5 and 6), but recombinant Med8C/18/20 could partially restore transcription (Figure 6, lanes 5 and 6). Thus, Med8C is essential for activated transcription in this assay. Since providing Med8C in *trans* can partially rescue the defect, Med8C apparently also interacts non-covalently with the core head, consistent with a reported two-hybrid interaction between Med8C and Med17 (Guglielmi et al, 2004). These functional data are consistent with the two-subcomplex architecture of the head (Figure 4A), and highlight the critical functional role of Med8C in tethering the two subcomplexes together.

#### **3.1.4 Med8C/18/20 is a functional submodule *in vivo***

To test whether the 'two-subcomplex architecture' of the Mediator head underlies Mediator function *in vivo*, and to determine whether the only role of Med8C is to tether Med18/20, we carried out comparative gene expression profiling with different yeast deletion strains. If the only role of Med8C is tethering of Med18/20 to Mediator, then the changes in gene expression observed in a Med8C deletion should be essentially the same as for deletion of either Med18 or Med20. We first used three strains of the same genetic background but specifically lacking one component of the Med8C/18/20 subcomplex. The expression profiles for these *med8C* $\Delta$ , *med18* $\Delta$ , and *med20* $\Delta$  deletion strains showed virtually the same pattern of changes in mRNA levels (Figure 7A). This is reflected in very high pairwise overall correlations (Figure 7B). For comparison, we repeated the analysis with two deletion strains that lacked genes for nonessential subunits Med2 or Med3 that reside in the tail module and also play a positive role in transcription (van de Peppel et al, 2005). Expression changes induced by Med2 or Med3 deletion correlate weakly with those induced by deletion of Med8C/18/20 components, although there is some overlap (Figure 7A and B). Therefore, the Med8C/18/20 subcomplex regulates transcription of a specific subset of genes and forms a functional submodule *in vivo*. The similarity in expression profiles is in agreement with the idea that the only role of Med8C is to tether Med18/20 to the core head. These results highlight the possibility to correlate structural data with transcriptome profiles, thereby identifying proteins and protein domains that reside in the same functional module.



**Figure 7: Comparative gene expression profiling**

(A) Diagram of genes exhibiting significantly altered mRNA levels (vertical axis) for different Mediator deletion strains (horizontal axis), clustered alongside the *med18Δ*, *med20Δ*, and *med8cΔ* expression profiles (indicated by black bar). Changes in mRNA levels compared with the wild-type strain are depicted in red (up), green (down), or black (no change). (B) Pearson's correlation matrix for expression profiles of strains *med8cΔ*, *med18Δ*, and *med20Δ*, and tail subunits deletion strains *med2Δ* and *med3Δ*. (C) Number of significantly altered genes in all three deletion strains of the Med8C/18/20 submodule. Genes not annotated in the GO database are depicted in gray, up-regulated genes are indicated in red, and down-regulated genes are shown in green. (D) Percentage of genes for which expression was significantly changed in all deletion mutants of the Med8C/18/20 (blue) submodule or Med2 and Med3 (yellow) compared with the percentage of the genome (black). All 32 biological processes from GO Slim Mapper are shown as follows: amino acid derivative metabolic process (1), anatomical structure morphogenesis (2), carbohydrate metabolic process (3), cell budding (4), cell cycle (5), cell wall organization and biogenesis (6), cellular homeostasis (7), cellular respiration (8), conjugation (9), cytokinesis (10), cytoskeleton organization and biogenesis (11), DNA metabolic process (12), electron transport (13), generation of precursor metabolites and energy (14), lipid metabolic process (15), meiosis (16), membrane organization and biogenesis (17), nuclear organization and biogenesis (18), organelle organization and biogenesis (19), protein catabolic process (20), protein modification process (21), pseudohyphal growth (22), response to stress (23), ribosome biogenesis and assembly (24), RNA metabolic process (25), signal transduction (26), sporulation (27), transcription (28), translation (29), transport (30), vesicle-mediated transport (31), and vitamin metabolic process (32). Overrepresented processes are marked with a circle. (E) Percentage of genes involved in conjugation, sporulation, and vitamin metabolic process for which expression was significantly changed in all deletion mutants of the Med8C/18/20 submodule or Med2 and Med3. Red and green histograms correspond to up- and down-regulated genes, respectively. The dotted line represents the percentage of the genome involved in the respective process.

### 3.1.5 Med8C/18/20 is required for transcription of specific genes

Analysis of the expression profiles revealed that a total of 165 mRNA levels were significantly altered in all three deletion strains of the Med8C/18/20 submodule. Of those, 117 were down-regulated, and 48 were up-regulated, showing that the module predominantly acts as a positive factor in transcription, but can also act as a negative factor for certain genes. Of the down- and up-regulated genes, 44% and 37%, respectively, were not annotated in the gene ontology (GO) database (Figure 7C) (Ashburner et al, 2000). Since it was previously shown that Med8C/18/20 binds to the TATA box-binding protein (TBP) *in vitro* (Larivière et al, 2006), we were encouraged to unravel a molecular basis for Med8C/18/20 function by searching for reoccurring promoter elements or upstream motifs in the deregulated genes, and for common transcription factors known to regulate these genes. This search was, however, unsuccessful, suggesting a complex context-dependent mechanism of Med8C/18/20 function.

We next analyzed the cellular function of genes regulated by Med8C/18/20 and by Med2 or Med3 with the GO Slim Mapper tool. Many of the affected biological processes, including amino acid derivative metabolism, carbohydrate metabolic process, and vitamin metabolic process, were overrepresented in both the Med8C/18/20-regulated and Med2/3-regulated genes (Figure 7D), reflecting the partial overlap of the expression profiles (Figure 7A). However, some biological processes were overrepresented only among Med8C/18/20- or Med2/Med3-regulated genes. In particular, genes involved in conjugation were specifically Med8C/18/20-regulated, whereas genes involved in sporulation as well as cell wall organization and biogenesis were specifically Med2/Med3-regulated (Figure 7E). These findings are consistent with the previously reported involvement of Med18 in conjugation (Holstege et al, 1998) and with a large-scale functional genomics analysis of sporulation efficiency (Enyenih & Saunders, 2003).

### 3.1.6 Down-regulation of nonactivated genes and basal promoter activity

Down-regulation of conjugation genes in *med8CΔ*, *med18Δ*, and *med20Δ* deletion strains was surprising, since their transcription was expected to be repressed under our experimental conditions. However, the same phenomenon was observed for other genes that should be repressed under optimal growth conditions. The GAL4 gene is further down-regulated in the *med8CΔ*, *med18Δ*, and *med20Δ* deletion strains, but not

in the control strains *med2Δ* or *med3Δ*. Genes involved in the catabolism of serine and glycine are further down-regulated in all deletion strains, including *CHA1*, *GCV1*, *GCV2*, and *GCV3*. *CHA1* shows the most down-regulated transcript levels, although it is expected to be expressed only under nitrogen-limiting conditions. Thus, Med8C/18/20 is required for low levels of transcription of nonactivated genes.

Taken together, many genes that are repressed under our growth conditions were apparently transcribed at a low level in a Med8C/18/20-dependent manner. The requirement of Med8C/18/20 for low-level transcription of nonactivated genes is consistent with the observation that even low levels of transcription require Mediator (Kornberg, 2005; Takagi et al, 2006). Although an indirect effect cannot be ruled out, our data suggest that Med8C/18/20 is important for basal promoter activity, and are consistent with a possible TBP interaction *in vivo*. A role of Mediator in basal transcription may be further investigated by similar analyses of other Mediator submodules and transcriptome profiling of several mutant yeast strains under various growth conditions.

### 3.1.7 Discussion

The Mediator head module is essential for cell viability, and a temperature-sensitive point mutation in the head leads to a global defect in transcription (Holstege et al, 1998). Here, we show that the head module contains the distinct conserved Med8C/18/20 submodule that is not essential for viability and regulates only a subset of genes. Identification and characterization of the Med8C/18/20 submodule required a combination of structural biology, yeast genetics, biochemistry, and transcriptome analysis. Our results support the idea that the known Mediator modules head, middle, and tail contain distinct submodules with different functions that are involved in the regulation of different subsets of genes.

More generally, we demonstrate how structural and functional information obtained on the molecular level *in vitro* can be correlated with changes on the systems level *in vivo*. In particular, the structure-guided design of mutant yeast strains enabled a precise disruption of molecular interactions and their functional analysis *in vitro* and *in vivo*. This approach is superior to the generally used gene deletion analysis, which does not take into account the consequences of such deletions for native protein complex structures. Most proteins reside in complexes (Gavin et al, 2006; Gavin et al, 2002; Krogan et al, 2006), and gene deletion will often result in complex disintegration and

malfunction, and thus in complicated changes of the transcriptome, which renders correlations between the molecular and the systems level difficult or impossible. In contrast, the structure-based systems perturbation analysis conducted here reduces the complexity of differential gene expression patterns, and facilitates the dissection of transcriptional coregulatory complexes into distinct functional submodules. In the future, this approach may be used for a reliable analysis of gene regulatory molecular networks. In particular, structure-based perturbation of cooperative molecular interactions could elucidate combinatorial and context-dependent gene regulatory mechanisms on the system level.

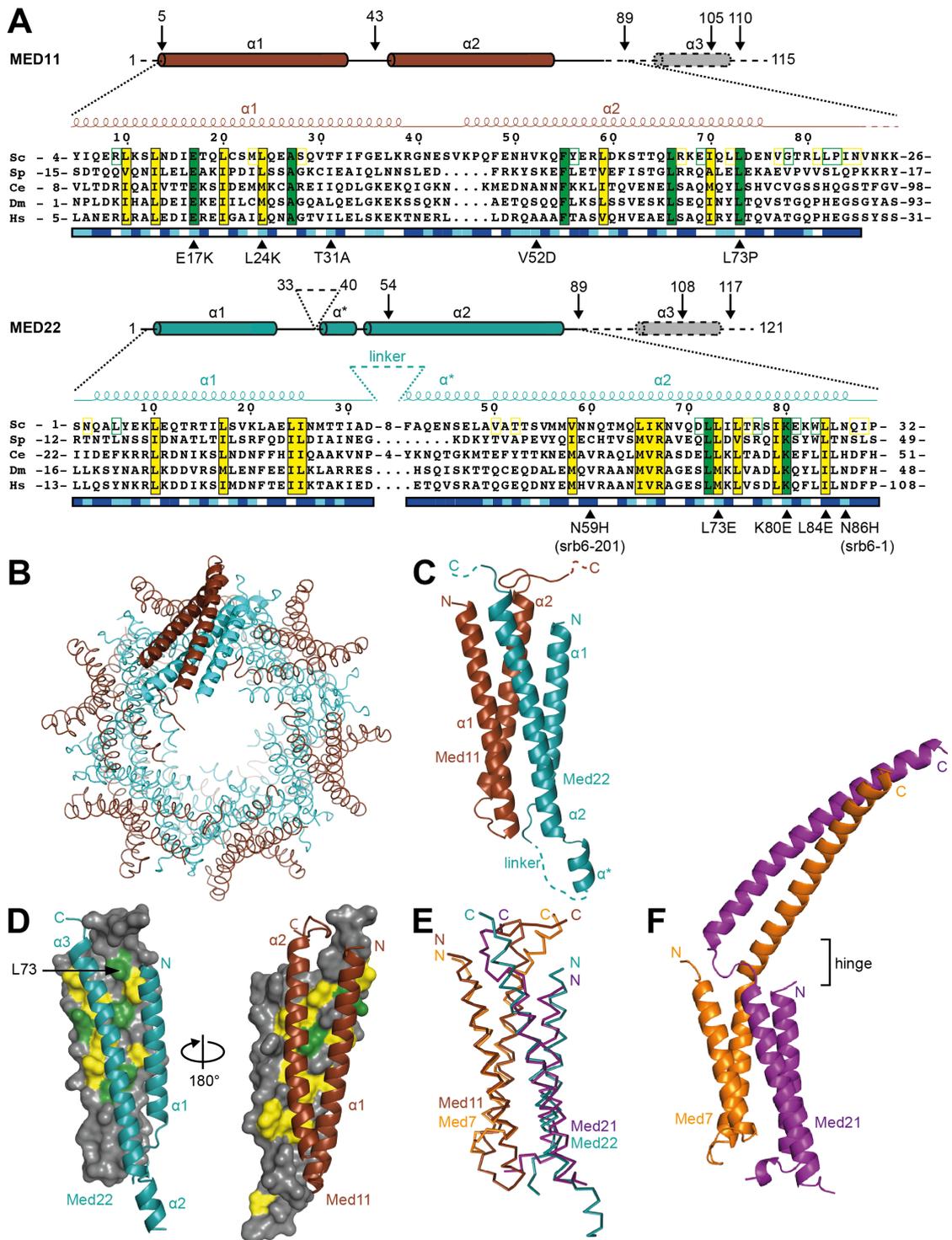
## **3.2 Mediator head subcomplex Med11/22 contains a common helix bundle building block with a specific function in transcription initiation complex stabilization**

### **3.2.1 Revised Med11 N-terminus**

Based on published interaction data (Esnault et al, 2008; Guglielmi et al, 2004; Takagi et al, 2006) we assumed that Sc Med11 and Med22 form a stable heterodimer. Indeed, the two subunits could be co-expressed in recombinant form (Larivière et al, 2006) in *E. coli* and subsequently co-purified. When we analyzed recombinant Med11/22 and endogenous purified Mediator by SDS-PAGE, we noticed a discrepancy in the electrophoretic mobility of Med11 (not shown). A comparison of the annotated Med11 open reading frame (ORF) with experimentally determined transcription start sites (Miura et al, 2006) revealed that the start codon of Med11 was incorrectly assigned in the databases. The real initiator methionine corresponds to residue 17 of the original annotation. We corrected the amino acid numbering and used the new numbering throughout.

### **3.2.2 Med11/22 structure solution**

We prepared the correct full-length recombinant Med11/22 heterodimer in pure recombinant form. In contrast to the initial preparation, this complex crystallized. The crystals could however not be refined, likely due to flexibility in the poorly conserved C-termini of both subunits. We therefore tested various truncated protein variants for their solubility and crystallization behavior, and found that a variant comprising the highly conserved core of Med11/22 (Med11<sub>5-89</sub>/Med22<sub>1-89</sub>, Figure 8A) resulted in crystals diffracting up to a resolution of 2.1 Å. The X-ray structure was solved by selenomethionine labeling and multi-wavelength anomalous diffraction, and was refined to a free R-factor of 20.8 % (Table 20). The crystals contained an unusual sphere-like arrangement of 12 heterodimers in the asymmetric unit (Figure 8B). Disordered in the crystals was only a non-conserved Med22 linker (residues 33-40) between helix  $\alpha$ 1 and the non-conserved helix  $\alpha^*$ . Deletion of this linker in yeast had no phenotype (Figure 9A).



**Figure 8: Structure of Med11/22 Mediator subcomplex**

(A) Multiple sequence alignment of the conserved core of Med11 and Med22 from *Saccharomyces cerevisiae* (Sc), *Schizosaccharomyces pombe* (Sp), *Caenorhabditis elegans* (Ce), *Drosophila melanogaster* (Dm) and *Homo sapiens* (Hs). The corrected numbering of Sc Med11 is used. Invariant and conserved residues are highlighted in green and yellow, respectively. Additionally, residues that are invariant or conserved among the yeast family Saccharomycotinae (*Sc*, *Candida glabrata*, *Candida albicans*, *Ashbya gossypii*, *Kluyveromyces lactis* and *Debaryomyces hansenii*) are highlighted with green and yellow frames on the Sc sequence, respectively. Surface accessibility is indicated below the sequences (blue, high; cyan, intermediate; white, buried). Secondary structure elements of the conserved

structural core are shown above the sequences (spirals,  $\alpha$ -helices; lines, ordered but without secondary structure; dashed lines, disordered in crystal construct). In addition schematic views of the full proteins are shown. Consensus secondary structure predictions (Cuff & Barton, 1999; Jones, 1999; Ouali & King, 2000) for the C-termini of Sc Med11 and Med22 are indicated with dashed lines and grey filling. All truncations relevant for this study are indicated with arrows and the residue number. Previously reported as well as mutations generated in this study are marked with black triangles. Sequence alignments were done with MUSCLE (Edgar, 2004) and figures were prepared with ESPript (Gouet et al, 1999). **(B)** Ribbon-model representation of the 12 heterodimers within the asymmetric unit. Med11 and Med22 are depicted in brown and cyan, respectively. **(C)** Ribbon-model representation of the Med11/22 crystal structure. Secondary structure elements are labeled according to (A). The linker between helices  $\alpha 1$  and  $\alpha^*$  of Med22 (residues 33-40) was disordered. C-termini of Med11 and Med22 had to be removed from crystallization construct and are therefore lacking in the structure. **(D)** Dimer interface conservation. On the *left*, Med11 is shown in surface representation together with a ribbon model of Med22. The view is related to (C) by a  $90^\circ$  rotation around a vertical axis. On the *right*, Med22 is shown in surface representation together with a ribbon model of Med11. The two views are related by a  $180^\circ$  rotation around a vertical axis. **(E)** Superimposition of  $C\alpha$  traces of the four-helix bundle folds of Med11/22 (brown and cyan; this study) and Med7C/21 (orange and purple; (Baumli et al, 2005)). **(F)** Ribbon-model representation of the complete Med7C/21 structure (pdb accession code 1YKE). The reported flexible hinge region is indicated.

**Table 20: Data collection and refinement statistics for Med11/22 structure**

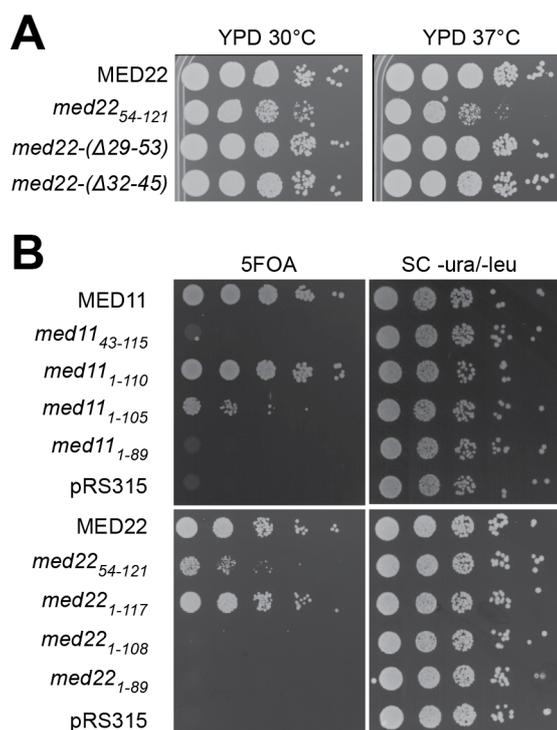
Factor	Peak	Inflection	Remote
Data collection			
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2		
Cell parameters			
a (Å)	148.5		
b (Å)	173.8		
c (Å)	101.6		
Wavelength (Å)	0.9796	0.9797	0.9720
Resolution range (Å) <sup>a</sup>	100-2.05 (2.10-2.05)	100-2.2 (2.26-2.20)	100-2.3 (2.36-2.30)
Completeness (%)	98.9 (99.1)	98.9 (99.1)	98.9 (98.7)
Unique reflections	315,134 (23,359)	256,151 (19,022)	224,540 (16,639)
Redundancy	2.9 (3.0)	2.9 (2.9)	2.9 (2.8)
R <sub>sym</sub> (%)	7.2 (57.8)	7.1 (52.6)	7.6 (55.9)
$\langle I \rangle / \langle s \rangle$	10.8 (2.0)	11.5 (2.2)	11.5 (2.2)
Refinement			
Number of residues	1,891		
Number of non-hydrogen atoms	16,435		
Number of solvent molecules	922		
RMS bond deviation (Å)	0.008		
RMS angle deviation (°)	1.006		
Ramachandran plot (preferred/allowed)	99.3/0.7		
R <sub>cryst</sub> (%)	17.0		
R <sub>free</sub> (%) <sup>b</sup>	20.8		

<sup>a</sup> The numbers in parenthesis correspond to the highest resolution shell

<sup>b</sup> 5% of the data were set aside for free R-factor calculation

### 3.2.3 A helix bundle building block in Mediator

Med11/22 forms an anti-parallel four-helix bundle (Figure 8C). This bundle fold is required *in vivo* since deletion of Med11 helix  $\alpha 1$  (*med11*<sub>43-115</sub>) was lethal in yeast, and deletion of  $\alpha 1$  and  $\alpha^*$  of Med22 (*med22*<sub>54-121</sub>) caused a growth defect (Figure 9B).



**Figure 9: *In vivo* phenotyping of Med11/22 truncations**

Yeast complementation assays. Med11 and Med22 constructs including 500 bp upstream of the start codon and 300 bp downstream of the stop codon were cloned into a pRS315 plasmid (LEU2) and transformed into the respective yeast shuffle strains. On 5FOA plates the URA3 shuffle plasmid encoding the respective full-length gene is shuffled out. Yeast cells lacking either the N- or C-terminus of Med11 or the C-terminus of Med22 are inviable. Cells lacking the N-terminal helix of Med22 or the last 10 amino acids of Med11 display a slow growth phenotype.

Unexpectedly, the bundle fold resembles the previously reported structure of the heterodimeric Med7C/21 subcomplex of the middle module (Baumli et al, 2005). The Med11/22 and Med7C/21 structures show a root mean square deviation of 3.1 Å over 121 C $\alpha$  atoms (Figure 8E). By combining results from HHPred (Söding et al, 2005) and secondary structure predictions, we detected a total of six possible heterodimeric four-helix bundles in Mediator. Nine out of 17 Mediator core subunits and two metazoan-specific subunits could participate in bundle formation (Figure 10). Heterodimer formation has been experimentally verified for Med11/22 in the head (this study), Med4/9 (Koschubs et al, 2010), and Med7/21 (Baumli et al, 2005) in the middle, and Med2/3 in the tail (Beve et al, 2005). Additional possible bundles include a Med10/14 heterodimer in the middle module (Koschubs et al, 2010) and a metazoan-specific Med28/30 subcomplex. These results establish the four-helix bundle fold as a common

building block of Mediator that is apparently found in all modules except for the dissociable kinase module.

		heterodimer formation	HHPred prediction Med7/21 bundle fold	
			p-value	score
Head module				
Med11		✓	0.0032*	22
Med22			0.0099	19.3
Middle module				
Med21		✓	N/A	N/A
Med7C			N/A	N/A
Med4		✓	0.00037	29.6
Med9			1.3E-05	35.9
Med10		✓	0.0068	21.4
Med14			-	-
Tail module				
Med2		✓	-	-
Med3			-	-
Metazoan only				
Med28		?	0.00032*	29.9
Med30			0.0018*	24.9

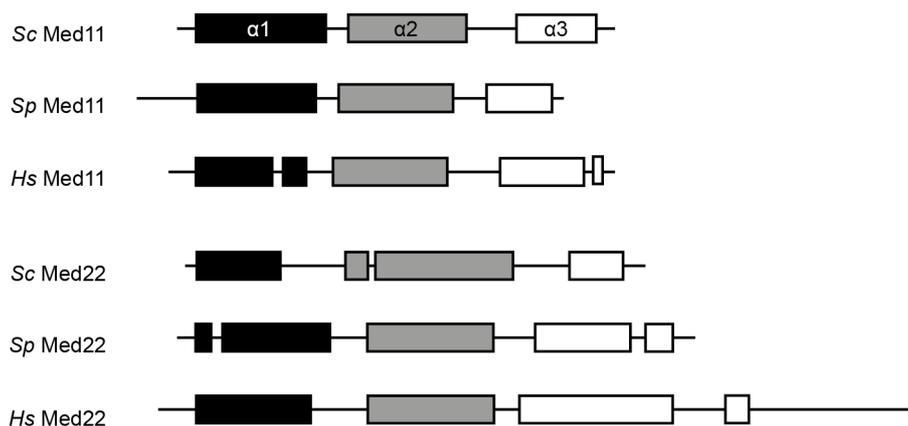
**Figure 10: A heterodimeric four-helix bundle building block in Mediator**

Schematic depiction of Mediator subunits predicted to share the heterodimeric bundle fold of Med11/22.  $\alpha$ -helices from crystal structures or from predictions (Cuff & Barton, 1999; Jones, 1999; Ouali & King, 2000) are indicated as boxes drawn to scale. Helix  $\alpha_1$ ,  $\alpha_2$  and C-terminal helical extensions are colored in black, grey and white, respectively. Check marks indicate experimentally confirmed heterodimers (co-expression and co-purification). Structural homology to the published Med7C/21 structure was predicted with HHPred (<http://toolkit.lmb.uni-muenchen.de/hhpred>) for all listed subunits except the highly divergent *Sc* subunits Med2 and Med3. The p-value and score for the HHPred searches are given. When the human protein sequence was used for the HHPred search, the p-values are marked with an asterisk.

### 3.2.4 Essential C-terminal helices extend from the Med11/22 helix bundle

In the structure of Med7C/21 (Baumli et al, 2005), the four-helix bundle is connected to C-terminal helical extensions from both subunits that form a coiled-coil (Figure 8F). The C-terminal extensions of Med11 and Med22 lack in the Med11/22 structure, but are predicted to each contain an additional helix (Figure 8A and Figure 11). However, in contrast to Med7C/21, the C-terminal helices are not predicted to form a coiled-coil, and are connected to the bundle through longer, probably flexible linkers. To

investigate the importance of the predicted Med11/22 C-terminal helices *in vivo*, we generated several Med11/22 truncation variants (Figure 8A) and tested them in yeast complementation assays (Figure 9A and B). Removal of either helix (*med11*<sub>1-89</sub> or *med22*<sub>1-89</sub>) or even partial truncation of the Med22 C-terminal helix (*med22*<sub>1-108</sub>) was lethal under standard conditions. Thus, the apparently flexible and helical C-terminal extensions from the Med11/22 bundle are required for normal cell growth.



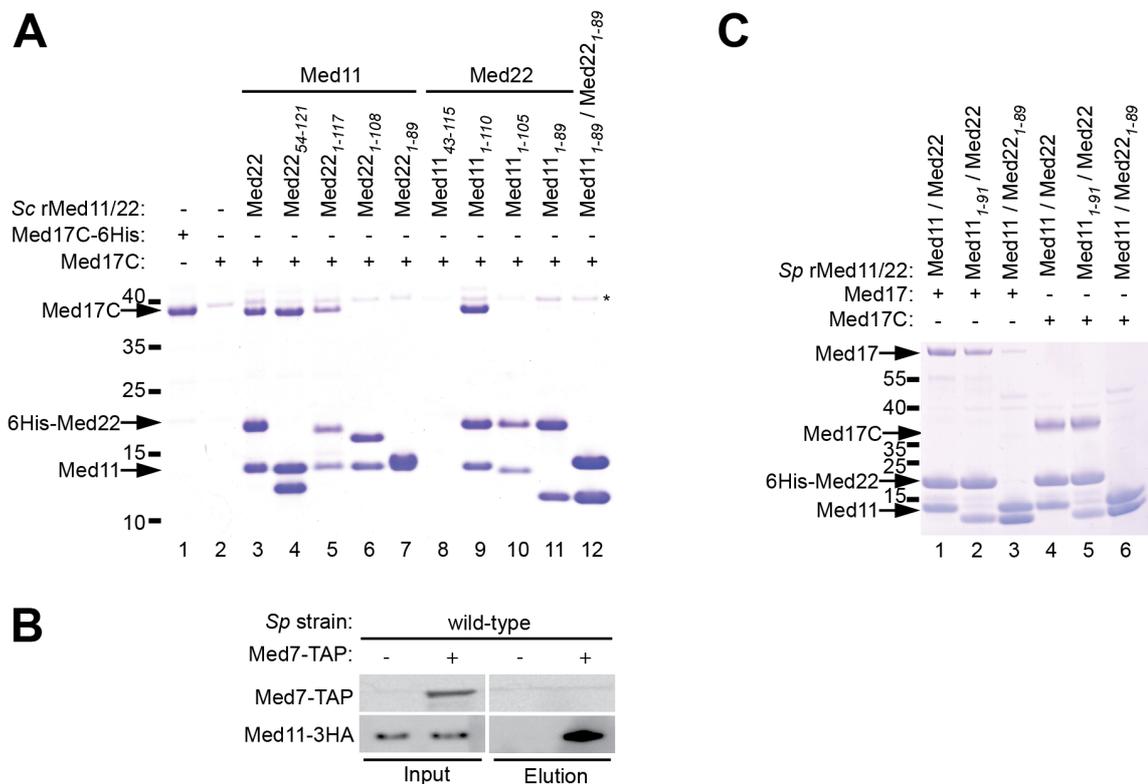
**Figure 11: C-terminal helices in Med11/22 are conserved across species**

Conservation of secondary structure elements in Med11 and Med22 across species (*Saccharomyces cerevisiae* (Sc), *Schizosaccharomyces pombe* (Sp) and *Homo sapiens* (Hs)). All secondary structure elements from crystal structures (when available) or from consensus secondary structure predictions are drawn to scale (boxes,  $\alpha$ -helices; lines, ordered but without secondary structure). Helix  $\alpha$ 1,  $\alpha$ 2 and C-terminal helical extensions are colored in black, grey and white, respectively.

### 3.2.5 Med11/22 extensions bind a Med17 C-terminal domain

We next asked whether the C-terminal extensions anchor the Med11/22 bundle within the head module by an interaction with Med17, the architectural head subunit. We could indeed detect *in vitro* binding of Med11/22 to a soluble C-terminal domain of Med17 (Med17C, residues 377-687) (Figure 12A, lane 1 and 3). To map the Med17 binding determinant in Med11/22, we co-expressed and co-purified truncated Med11/22 variants with Med17C in *E. coli*. Recombinant Med11/22 heterodimers containing Med11<sub>43-115</sub> could not be co-purified, whereas Med11<sub>1-89</sub>, Med11<sub>1-105</sub>, Med22<sub>54-121</sub>, and Med22<sub>1-89</sub> all formed stable complexes (Figure 12A). Truncation of either predicted C-terminal helix (Figure 12A, lane 6, 7, 10-12) prevented co-purification of Med17C with Med11/22, while truncation of the Med22 N-terminus (Figure 12A, lane 4) or C-terminal truncations that did not affect the C-terminal helices

(Figure 12A, lane 5 and 9) had no effect. We conclude that the C-terminal helices of Med11/22 anchor the four-helix bundle to Med17C, and that the C-terminus of Med22 is absolutely required for this interaction *in vitro* and *in vivo*. This is in accordance with yeast-two-hybrid results showing that the Sc Med11-Med17 interaction is lost upon C-terminal truncation of Med11 (Esnault et al, 2008). The consistency of these *in vitro* binding data with the above *in vivo* phenotyping further suggests that Med11/22 anchoring to Med17C is essential for cellular growth.



**Figure 12: C-terminal extensions of Med11/22 bind a Med17 C-terminal domain**

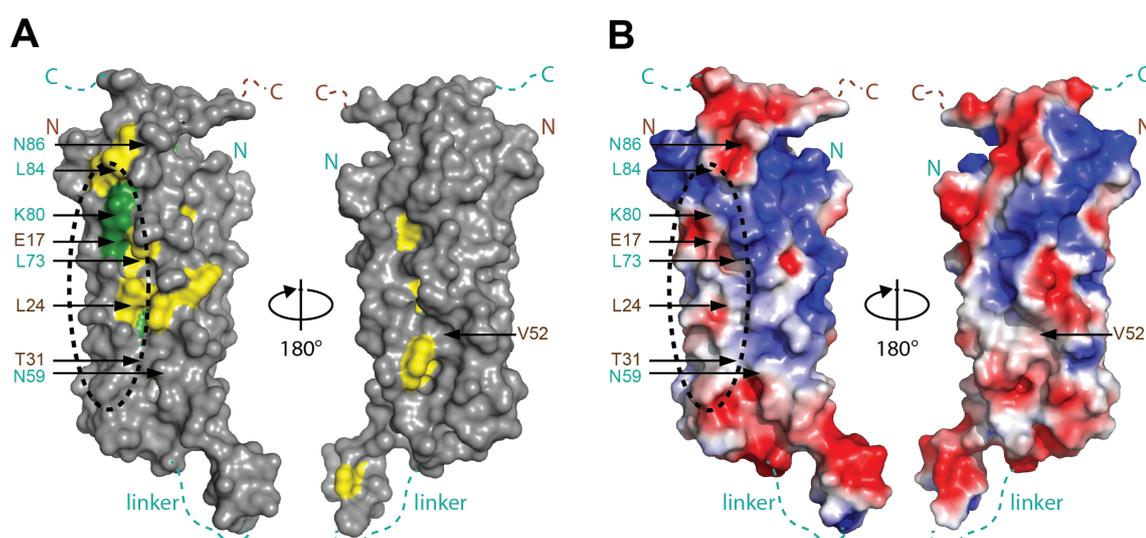
**(A)** Co-expression in *E. coli* and co-purification of Sc a C-terminal domain of Med17C (residues 377-687) with Sc 6His-Med22/Med11 constructs using nickel magnetic beads. A co-purifying contaminant is marked with an asterisk. **(B)** Co-immunoprecipitation of the putative *Sp* Med11-3HA with *Sp* Med7-TAP from *Sp* lysate using IgG agarose beads. **(C)** Co-expression in *E. coli* and co-purification of *Sp* Med17 (lanes 1-3) and *Sp* Med17C (residues 257-545; lanes 4-6) with *Sp* His6-Med22/Med11 constructs using nickel magnetic beads.

### 3.2.6 A conserved Med17C/11/22 Mediator subcomplex

The helix bundle of Med11/22 appears to be conserved among eukaryotes, because amino acid residues that constitute the Med11-Med22 interface are conserved (Figure 8D). However, a Med11 homologue has thus far not been found in *Sp* (Spahr et al, 2001). We could however predict a remote Med11 homology in the *Sp* protein SPAC644.10 (Wood et al, 2002). Indeed, *Sp* SPAC644.10 could be co-immunoprecipitated with Med7, suggesting it is a Mediator subunit (Figure 12B). Amplification and sequencing of the SPAC644.10 cDNA clone revealed the absence of an annotated intron, and enabled co-expression and co-purification of recombinant *Sp* SPAC644.10 with *Sp* Med22 (Figure 12C). This establishes SPAC644.10 as the *Sp* Med11 homologue and confirms the conservation of the Med11/22 bundle. To investigate whether the conservation extends to the anchoring of Med11/22 on Med17C, we tested whether a trimeric *Sp* Med17/11/22 subcomplex could be obtained after subunit co-expression in *E. coli*, and this was indeed achieved (Figure 12C, lane 1). We mapped a soluble C-terminal domain of *Sp* Med17 (Med17C, residues 257-545) that was sufficient for Med11/22 binding (Figure 12C, lane 4). Truncation of the C-terminus of *Sp* Med22 did not affect Med11/22 heterodimer formation, but abolished Med17 binding (Figure 12C, lane 3 and 6). Truncation of the C-terminus of *Sp* Med11 (Med11<sub>1-91</sub>) had almost no effect (Figure 12C, lane 2 and 5). Thus, the interaction of the C-terminal extension from the Med11/22 bundle with Med17C is conserved between the distantly related yeast species *Sp* and *Sc*, and the Med17C/11/22 subcomplex is therefore a conserved architectural unit of the Mediator head.

### 3.2.7 A highly conserved interaction patch on Med11/22

Despite the importance of the head module for PIC formation, only few contacts between PIC components and head module subunits have been reported. Thus, we wanted to identify potential interaction surfaces on Med11/22 that could account for contacts to PIC components. Plotting sequence conservation (Figure 13A) and electrostatic surface charge (Figure 13B) onto the Med11/22 structure revealed a large, highly conserved surface patch with exposed hydrophobic residues on one side of the bundle domain. Several known Med11/22 mutations (Esnault et al, 2008; Lee et al, 1998; Thompson et al, 1993), namely *srb6-201* (*med22-N59H*), *srb6-1* (*med22-N86H*), *med11-T31A*, are located around this patch (Figure 13A and B).



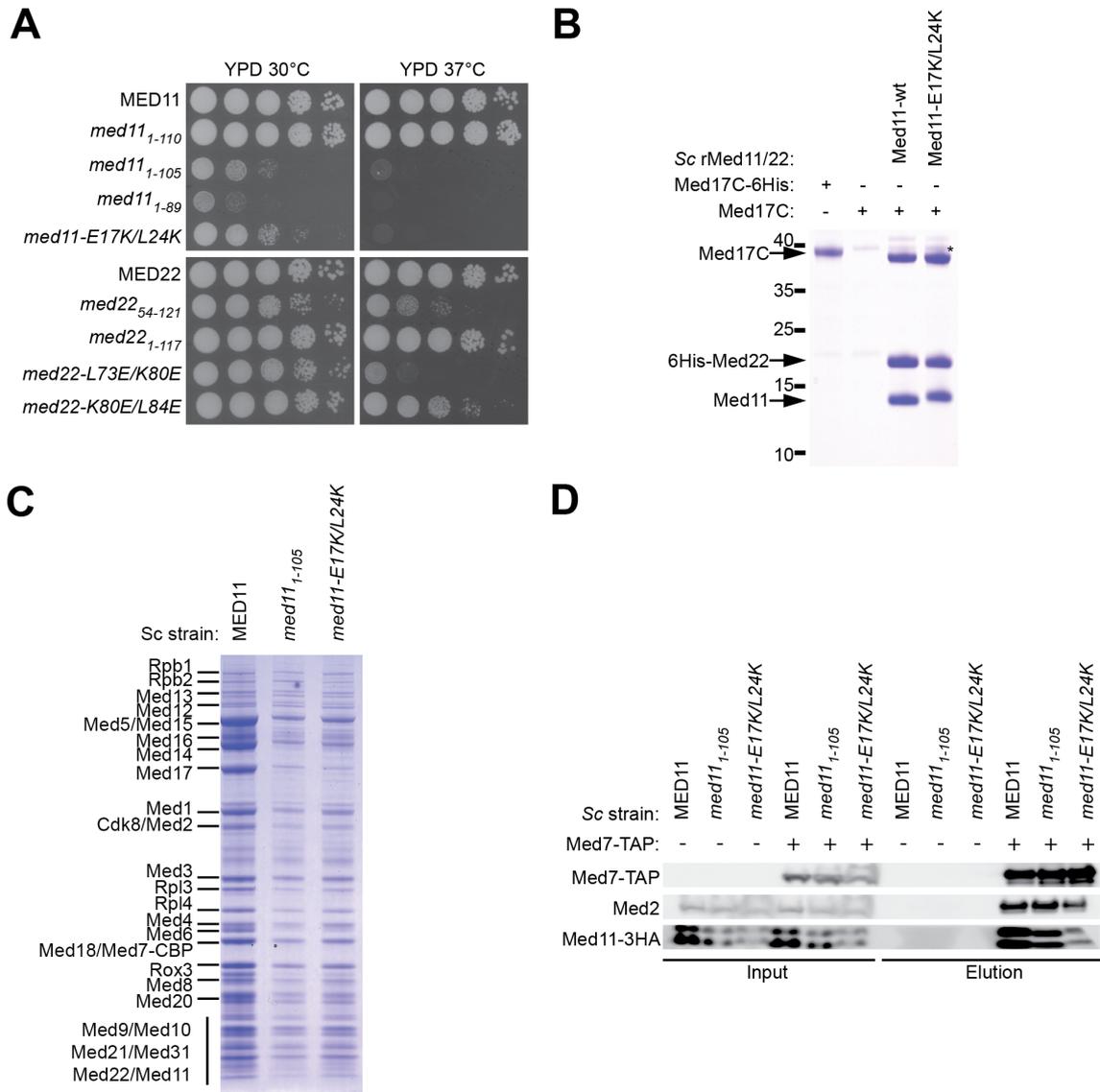
**Figure 13: Surface of Med11/22 helix bundle.**

**(A)** Surface conservation of the Med11/22 four-helix bundle. Invariant and conserved residues from yeast to human are highlighted in green and yellow, respectively. Residues of Med11 (brown) and Med22 (cyan) targeted by mutagenesis in this study and in previous reports are indicated with arrows. The orientation is identical to Figure 8C. The two views are related by a 180° rotation around the vertical axis. **(B)** Surface charge distribution. Red, blue and white areas indicate negative, positive and neutral charges, respectively. The two views and labeled residues are the same as in (A).

Mutations reported to affect interaction of Med11 with Med22 or Med17 (Esnault et al, 2008; Han et al, 1999), namely *med11-V52D*, *med11-G92S* and *med11-L73P*, are however located distant from this patch or in the hydrophobic core (Figure 13A and Figure 8D). This suggests that the conserved surface patch is involved in the interaction with components of the PIC.

To investigate the functional importance of this patch we generated structure-based surface mutations. Strains carrying double point mutations of adjacent surface residues (*med11-E17K/L24K*, *med22-L73E/K80E* and *med22-K80E/L84E*) exhibited various degrees of temperature-sensitivity (Figure 14A). The *med11-E17K/L24K* strain exhibited a strong growth defect at all temperatures, similar to *med11<sub>1-105</sub>*. However, unlike the Med11 truncation, the patch mutation did not affect anchoring of Med11/22 to Med17C *in vitro* (Figure 14B). Since Med11/22 is essential for Med17 stability (Takagi et al, 2006) we purified Mediator from MED11, *med11<sub>1-105</sub>* and *med11-E17K/L24K* yeast strains. Mediator composition and integrity appeared unchanged on a Coomassie-stained SDS-PAGE (Figure 14C) and was additionally confirmed by co-immunoprecipitation and Western Blot (Figure 14D). Thus the

phenotype observed for the Med11 mutants is not the result of impaired Mediator integrity, but apparently a direct effect of impaired Med11/22 interactions with other factors.

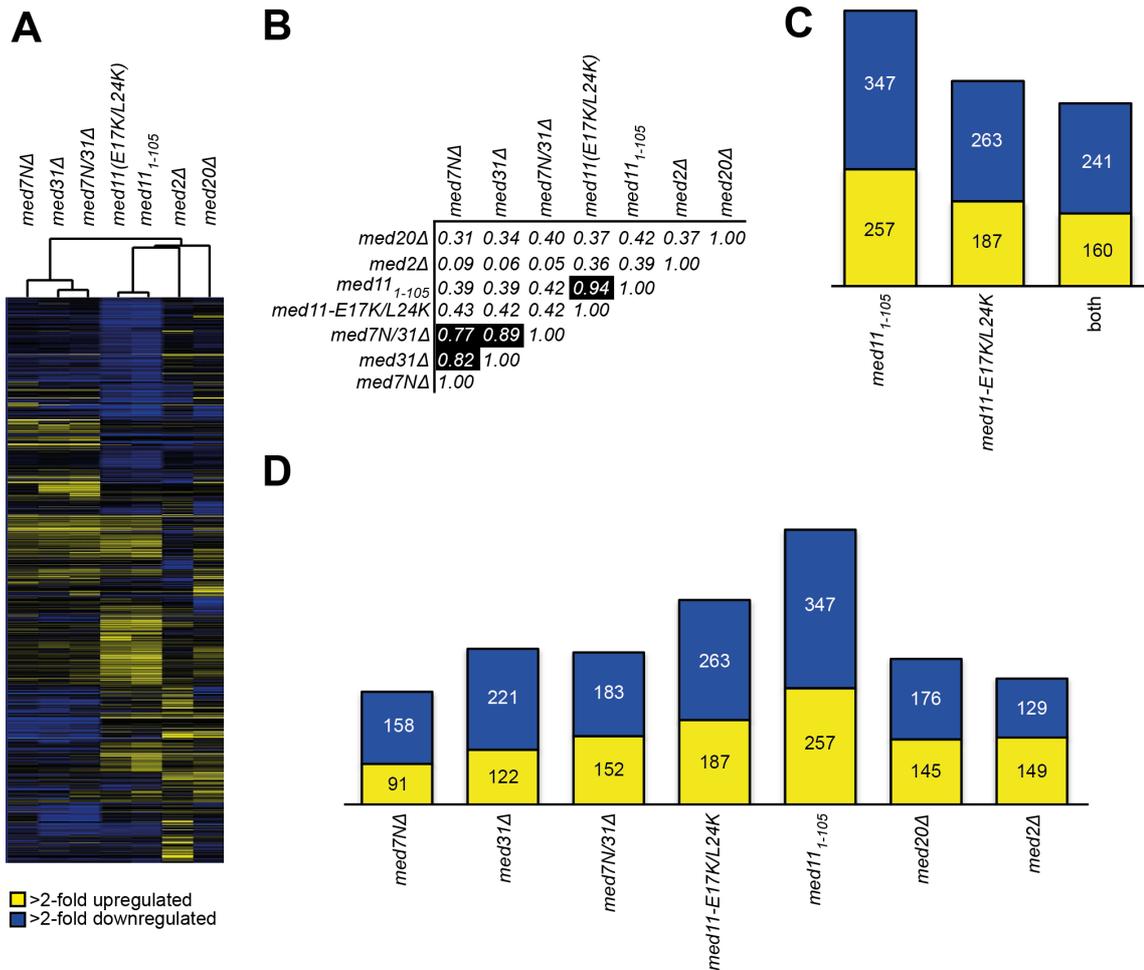


**Figure 14: A conserved interaction patch on Med11/22**

**(A)** Spot dilutions of yeast strains carrying structure-based mutations on the conserved surface patch and viable truncations (Figure 2) on YPD plates at 30°C and 37°C. **(B)** Co-expression in *E. coli* and co-purification of Sc Med17C with Sc 6His-Med22/Med11 and Sc 6His-Med22/Med11-E17K/L24K using nickel magnetic beads. **(C)** Mediator tandem affinity purification through Med7-TAP from wild-type, *med11*<sub>1-105</sub> and *med11-E17K/L24K* yeast strains. Co-purifying proteins were separated on a 5-12% gradient gel (Invitrogen) and bands were stained with Coomassie blue. Mediator subunits and common co-purifying contaminants are labeled. Med7-CBP marks Med7 after tandem-affinity purification still carrying the calmodulin binding protein-tag but lacking the cleaved protein A-tag. **(D)** Co-immunoprecipitation of Sc Med11-3HA (Mediator head module) and Sc Med2 (Mediator tail module) with Sc Med7-TAP (Mediator middle module) from wild-type, *med11*<sub>1-105</sub> and *med11-E17K/L24K* yeast strains.

### 3.2.8 Med11/22 is a functionally distinct submodule

To test whether mutations in Med11/22 affect transcription *in vivo*, we performed genome-wide gene-expression profiling of the *med11*<sub>1-105</sub> and *med11-E17K/L24K* yeast strains (Figure 15A).



**Figure 15: Med11/22 is a functional submodule regulating a specific subset of genes**

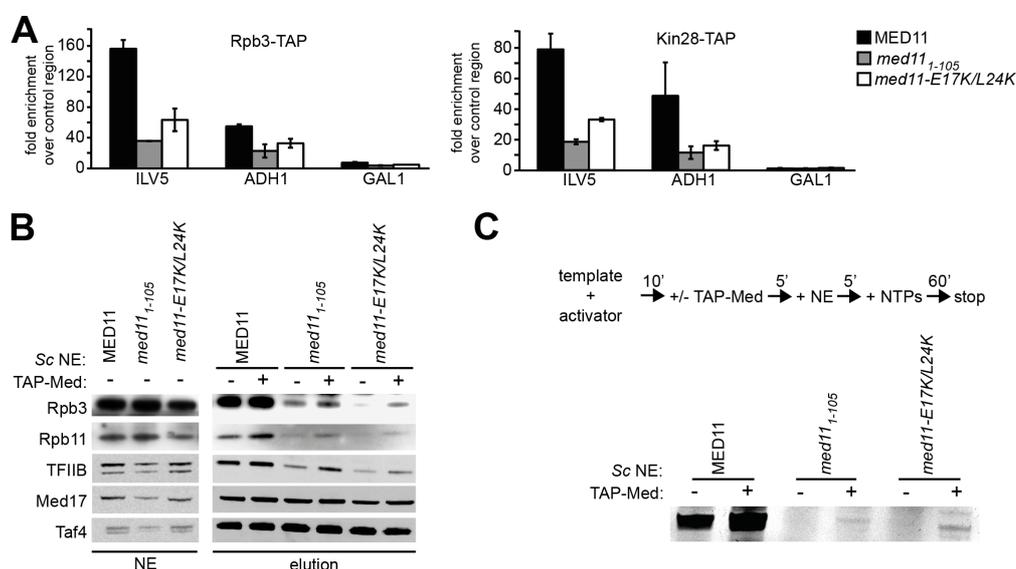
**(A)** Hierarchical cluster diagram (Pearson correlation) of genes exhibiting significantly altered mRNA levels (at least 2-fold, vertical axis) for different Mediator mutant strains (horizontal axis). Changes in mRNA levels compared to the wild-type strain are depicted in yellow (up), blue (down), or black (no change). **(B)** Pearson correlation matrix for expression profiles of different Mediator mutants (1, very high correlation; 0, no correlation; -1, very high anti-correlation). **(C)** Number of genes significantly (fold change > 2.0,  $p$ -value < 0.05) up-regulated (yellow) and down-regulated (blue) in *med11*<sub>1-105</sub>, *med11-E17K/L24K* and in both strains. **(D)** Comparison of gene expression changes of all mutants analyzed in this study. Number of genes significantly (fold change > 2.0,  $p$ -value < 0.05) up-regulated (yellow) and down-regulated (blue) is shown.

The Med11 mutants form a distinct cluster with a Pearson correlation coefficient of 0.94, correlating only moderately with *med20* $\Delta$  (head module), *med2* $\Delta$  (tail module) and with mutations affecting the previously described functional submodule Med7N/31 (middle module) (Koschubs et al, 2009) (Figure 15B). In total, 604 and 450 genes were significantly changed (fold expression change > 2-fold; p-value < 0.05) in *med11*<sub>1-105</sub> and *med11-E17K/L24K*, respectively. A total of 401 genes were significantly changed in both mutants, with 160 genes up-regulated and 241 genes down-regulated (Figure 15C). The number of significantly changed genes in both Med11 mutant strains is higher than in any other Mediator mutant tested (Figure 15D). The effect on gene expression is rather pleiotropic without any significant enrichment for specific Gene Ontology terms (Boyle et al, 2004). This indicates a distinct function of the Med11/22 subcomplex in gene regulation, and a more global role in contrast to the more gene-specific roles of previously reported non-essential Mediator submodules (Koschubs et al, 2009; Larivière et al, 2008; van de Peppel et al, 2005).

### 3.2.9 The Med11/22 surface patch functions in PIC stabilization

Since Med11 was previously reported to stabilize the TFIID subcomplex of TFIIF and, to some extent, Pol II at promoters (Esnault et al, 2008), we performed chromatin immunoprecipitations of Kin28, the kinase subunit of TFIID, and the Pol II subunit Rpb3 in wild-type, *med11*<sub>1-105</sub>, and *med11-E17K/L24K* strains (Figure 16A). In the mutant strains, we observed decreased occupancies for both Kin28 and Rpb3 at active promoters, indicating a defect in stable PIC formation. To test whether the decrease is a direct effect of impaired Med11/22 function or an indirect effect of global deregulation of gene expression, we performed *in vitro* immobilized template assays with yeast nuclear extracts (Figure 16B). PICs were assembled on an immobilized HIS4 yeast promoter in a Gcn4-dependent manner, washed and subsequently analyzed by Western blot. Consistent with the *in vivo* results, both mutant extracts displayed a decreased occupancy of Pol II after washing. As expected, the occupancy of TFIIB, interacting directly with Pol II, is also decreased. Mediator and TFIID, which are recruited directly by the transcriptional activator Gcn4 (Herbig et al, 2010), remain unaffected. Consistently, nuclear extracts from the Med11 mutant strains were inactive in *in vitro* transcription assays (Figure 16C). Since addition of purified wild-type Mediator to the mutant extracts partially rescued the recruitment of the basal machinery and consequently also transcriptional activity, the observed defects can be directly

linked to Mediator function. These results show that the conserved surface patch on the Med11/22 bundle is required for stable PIC formation.



**Figure 16: Med11/22 is required for stable PIC formation *in vitro* and *in vivo***

**(A)** Chromatin immunoprecipitation of Rpb3-TAP (Pol II) and Kin28-TAP (TFIIH kinase module) in wild-type, *med11<sub>1-105</sub>* and *med11-E17K/L24K* strains grown to early exponential phase in YP medium containing 2% glucose. Fold enrichment over a heterochromatic control region is shown for the yeast promoters of a highly expressed gene (ILV5), a housekeeping gene (ADH1) and a glucose-repressed gene (GAL1). **(B)** *In vitro* PIC assembly of wild-type, *med11<sub>1-105</sub>* and *med11-E17K/L24K* nuclear extracts (NE) on the immobilized HIS4 yeast promoter. PIC formation of mutant extracts was partially rescued by adding 2 pmol tandem-affinity purified Mediator complex (TAP-Med) to the nuclear extracts prior to PIC assembly. Presence of Pol II (Rpb3 & Rpb11), TFIIIB, Mediator (Med17) and TFIID (Taf4) was tested by Western Blot. **(C)** *In vitro* transcription assay with wild-type, *med11<sub>1-105</sub>* and *med11-E17K/L24K* nuclear extracts (NE) on the HIS4 yeast promoter. Order of addition is shown on top. Transcription was partially rescued by adding 2 pmol tandem-affinity purified Mediator complex (TAP-Med).

### 3.2.10 Discussion

To understand the regulatory mechanisms of eukaryotic transcription, a detailed structural and functional dissection of the involved multiprotein complexes is required. In this study, we extend our previous structure-function analysis of the general transcription coactivator Mediator. We combine structural biology, yeast genetics, biochemistry, and gene-expression profiling, to define and functionally characterize the Mediator subcomplex Med11/22, to provide a more detailed understanding of Mediator head module architecture, and to obtain insights into Mediator conservation and evolution. We report the structure of a conserved four-helix bundle domain in Med11/22 that puts reported mutations (Esnault et al, 2008; Han et al, 1999) into a molecular

framework and identifies a highly conserved surface patch that is required for a distinct widespread function of the Med11/22 subcomplex.

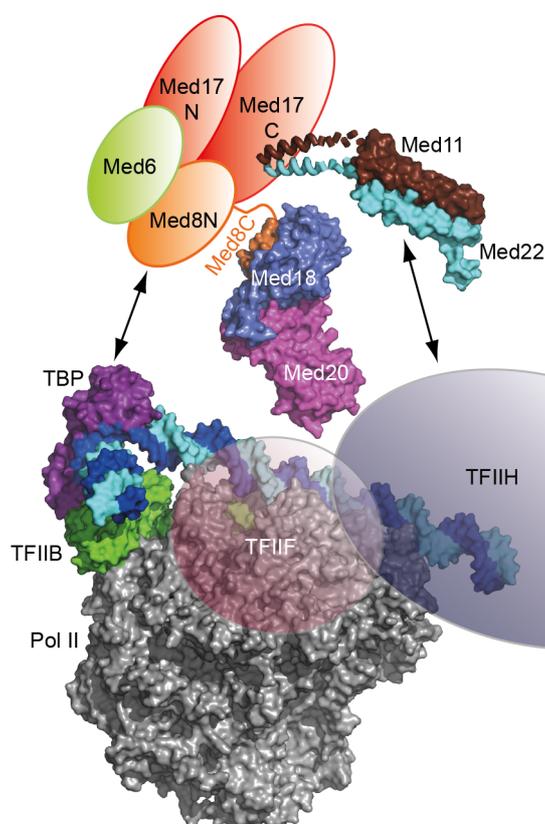
We suggest that the Med11/22 surface patch mediates an essential interaction with TFIIH, because the previously described temperature-sensitive mutation *med11-T31A* (Esnault et al, 2008) locates near the patch. This mutation was shown to decrease interaction of Med11 with the TFIIH subunit Rad3 in yeast-two hybrid assays, and promoter occupancy of the TFIIH kinase module TFIIK. In contrast to the mutation *med11-T31A*, which was discovered genetically, the mutation *med11-E17K/L24K* reported here was identified based on structural data, and had a strong general growth defect that allowed us to conduct functional studies under standard growth conditions. Chromatin immunoprecipitation showed a decrease of TFIIK and Pol II occupancies at active promoters *in vivo*, indicating a defect in PIC formation. Yeast nuclear extracts from the mutant strain were defective in stable PIC formation and inactive in transcription *in vitro*. The defects result from impairing the function of the Med11/22 submodule, and not the head module *per se*, since, in contrast to the reported Med17 mutant *srb4-138* (Ranish et al, 1999), activator-mediated recruitment of Mediator to the promoter was not affected. Thus the Med11/22 submodule specifically functions in promoting stable PIC formation. We propose that the conserved submodular architecture of the Mediator head enables multiple transient interactions with PIC components, including TBP (Kang et al, 2001) and TFIIH (Esnault et al, 2008) (Figure 7), thereby stabilizing the PIC and facilitating open complex formation and initial RNA synthesis.

Another intriguing observation is the presence of up to six related helix bundle folds, to which about half of the Mediator core subunits contribute. The presence of related bundle domains in different Mediator modules elucidates the evolutionary origin of Mediator. Large protein complexes with high functional modularity might generally have evolved through duplication of genes encoding dimers (Pereira-Leal et al, 2007). For example, the 15-subunit TFIID complex contains five heterodimeric subcomplexes interacting through a common histone-fold domain (Gangloff et al, 2001; Leurent et al, 2002). Subsequent divergent evolution of paralogous subunits could then generate asymmetry and diversification of protein interactions for functional specialization. Rapid evolution of Mediator by dimer duplication and diversification could explain how it is possible that the general transcription machineries are conserved between archaea and eukaryotes, whereas Mediator is only present in eukaryotes.

### 3.3 Mediator head module controls preinitiation complex formation

#### 3.3.1 A recombinant head module is functionally active

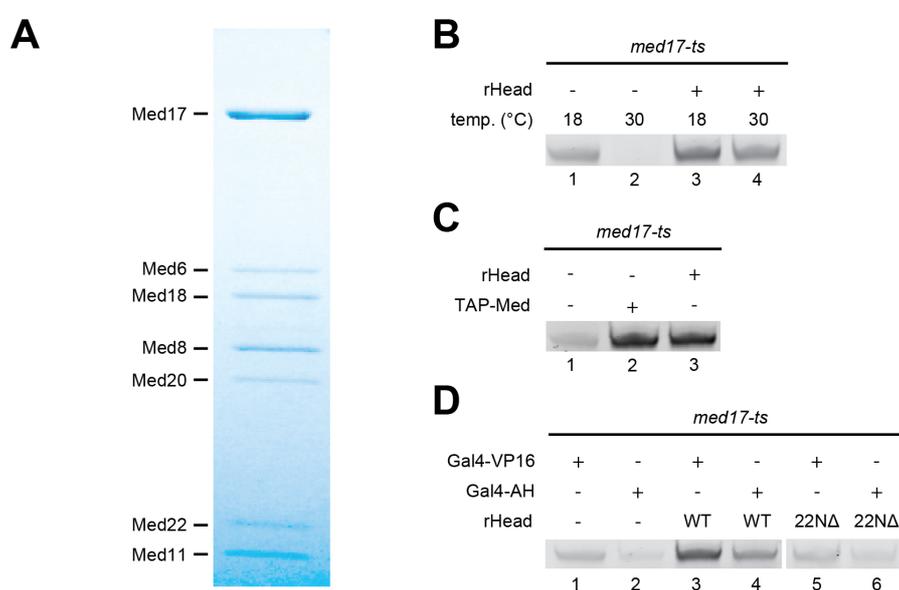
In our previous studies we had used structure-function analyses to demonstrate the existence of two functionally distinct submodules, Med8C/18/20 (Chapter 3.1) and Med11/22 (Chapter 3.2), within the *Sc* Mediator head module (Figure 14). To further characterize the architecture of the Mediator head module and its essential role during PIC assembly, structural and functional information on the complete head module is required. Therefore, the recombinant coexpression in *E. coli* and subsequent copurification of all seven *Sc* head subunits was established (Figure 18A; unpublished results from the Cramer laboratory). A similar approach using multicistronic expression plasmids had already been established for the 7-subunit *Sc* Mediator middle module (Koschubs et al, 2010).



**Figure 17: Model of the submodular architecture of Mediator and PIC contacts**

Crystal structures of the essential Med11/22 four-helix bundle (this work), the previously described non-essential Med8C/18/20 subcomplex (Larivière et al, 2006), and a molecular model of the Pol II–TBP–TFIIB–DNA promoter closed complex (Kostrewa et al, 2009) are drawn to scale. Locations of the general factors TFIIF and –H, as determined by biochemical probing (Chen et al, 2007; Kim et al, 2000), are indicated by semi-transparent ellipsoids. Mediator head interactions with PIC components, namely Med8/18/20 – TBP (Larivière et al, 2006) and Med11/22 – TFIIH (Esnault et al, 2008), are indicated by arrows.

To test the biological activity of the recombinant head module, we performed *in vitro* transcription assays in a nuclear extract from the temperature-sensitive yeast strain *med17-ts* (*srb4-138* strain). Whole cell extracts from this strain had been previously shown to be impaired in basal and activated transcription *in vitro* (Takagi & Kornberg, 2006). Indeed, *in vitro* transcription in *med17-ts* nuclear extracts was specifically impaired at the non-permissive temperature (Figure 18B, lanes 1 and 2). Addition of recombinant 7-subunit head module rescued the defect in activated transcription (Figure 18B, lanes 3 and 4). Addition of TAP-purified wild-type Mediator rescued the defect to a similar extent (Figure 18C).



### Figure 18: Recombinant Sc Mediator head module is functionally active

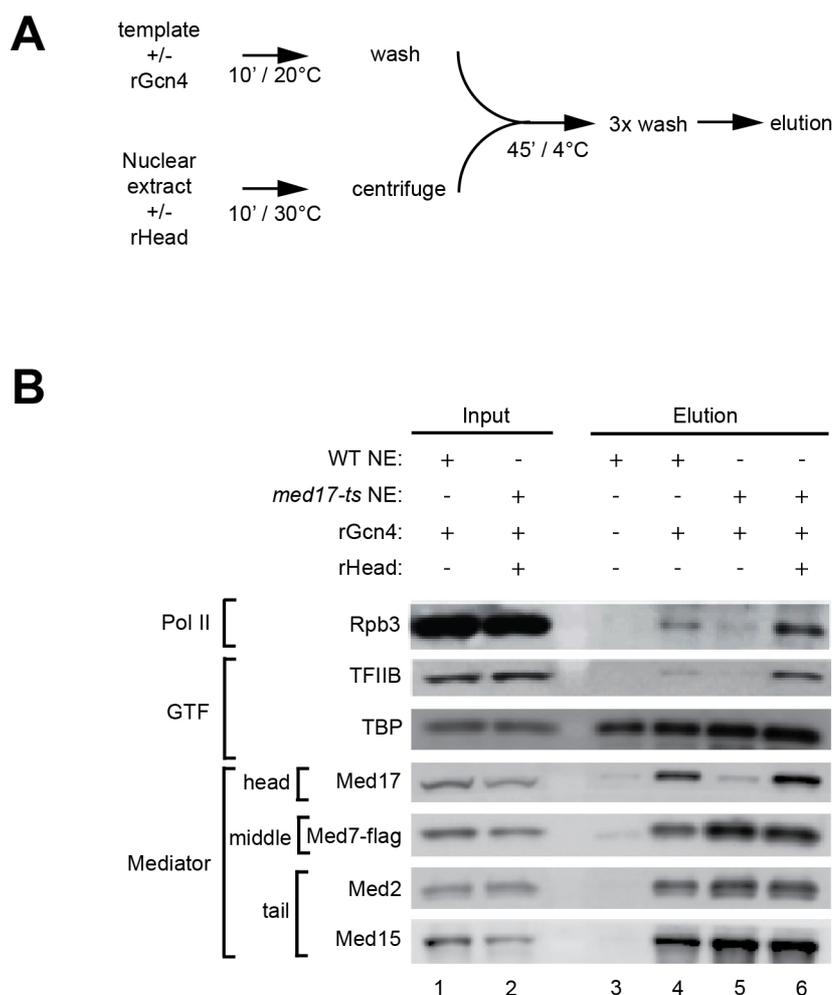
**(A)** Recombinant Sc Mediator head module preparation. All seven subunit were coexpressed in *E. coli* and subsequently copurified to homogeneity. **(B)** *In vitro* transcription in a temperature-sensitive nuclear extract prepared from the *srb4-138* yeast strain (*med17-ts*). The recombinant fusion activator Gal4-VP16 was added to all samples to monitor activated transcription. *In vitro* transcription was performed either at the permissive temperature (18°C; lane 1 and 3) or the non-permissive temperature (30°C; lane 2 and 4). Defect in activated transcription was rescued by addition of 2 pmol recombinant head (rHead). **(C)** *In vitro* transcription at 30°C as in (A). Addition of approximately 2 pmol TAP-purified Mediator (lane 2) rescues to transcription to a similar extent as addition of 2 pmol rHead (lane 3). **(D)** *In vitro* transcription at 30°C as in (A). Additionally Gal4-AH activated transcription was tested. Only addition of 2 pmol recombinant wild-type head module (WT; lane 3 and 4) rescued transcription while addition of 2 pmol recombinant head module containing a N-terminally truncated Med22 variant (22NΔ; residues 54-121) did not.

Furthermore, the rescue of activated transcription *in vitro* was observed with different activation domains (Figure 18D, lanes 1-4) and relied on the functional activity of the recombinant head module (Figure 18D, lanes 5 and 6). Similar to a *med22 $\Delta$*  yeast nuclear extract, which was unable to support activated transcription *in vitro* (Chapter 3.2), addition of a recombinant head module lacking the N-terminus of Med22 did not rescue activated transcription (Figure 18D, lane 5 and 6). These results demonstrate that the recombinant 7-subunit Sc Mediator head module preparation is biologically active and can be used for further biochemical and structural studies.

### 3.3.2 Preinitiation complex formation requires the head module

Since activated transcription in a *med17-ts* nuclear extract depends on a functional Mediator head module *in vitro*, we wanted to determine the factors specifically recruited by the Mediator head module. Previous studies had suggested that PIC formation as well as Mediator recruitment is abolished in a *med17-ts* nuclear extract (Ranish et al, 1999). This makes a distinction between complete Mediator and Mediator head module function impossible. However, due to the lack of a recombinant head module for add-back experiments, secondary effects from extract preparation could not be ruled out. Furthermore, the immobilized template assays had been performed with an artificial fusion template comprising a modified HIS4 core promoter but lacking the native upstream activating sequence. We hypothesized that the upstream activating sequence and the respective bound activators could influence the stability of Mediator and the basal transcription machinery at the core promoter. Therefore, we performed immobilized template assays with yeast nuclear extracts on a native HIS4 promoter template. Since the HIS4 UAS comprises several Gcn4 binding sites, we used recombinant full-length Gcn4 for activator-dependent PIC assembly. Consistent with the *in vitro* transcription assays (Figure 18B) we performed a 30°C heat step with all extracts prior to PIC assembly (Figure 19A). In a wild-type nuclear extract Mediator and PIC components were recruited to this native promoter template in a Gcn4 dependent manner (Figure 19B, lanes 3 and 4). Consistent with previous reports, TBP was recruited independent of the activator through a direct interaction with the TATA box (Ranish et al, 1999). Surprisingly and in contrast to previous studies, recruitment of Mediator middle and tail modules (Figure 19B, lanes 5) was not affected in the *med17-ts* nuclear extract. Only the recruitment of the head module (Med17) and PIC formation (Pol II and TFIIB) was abolished. Addition of recombinant 7-subunit head module

restored PIC formation (Figure 19B, lanes 6) consistent with the observed rescue of activated transcription *in vitro* (Figure 18). These results suggest that a stable head-less Mediator (termed Mediator body), comprising the middle and tail module but lacking the head module, can be recruited to the promoter by gene specific activators. However, only when the head module is bound, a stable PIC is formed.

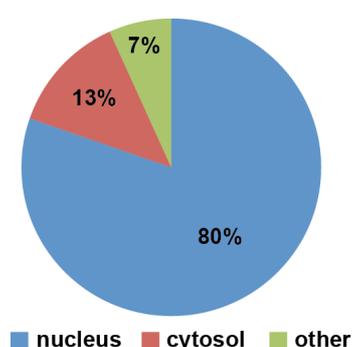


**Figure 19: PIC formation requires the Mediator Head module**

**(A)** Order of addition and incubation steps for immobilized template assay. Heat inactivation of temperature-sensitive nuclear extracts is done for 10 min at 30°C in the presence of rHead where applicable. **(B)** Immobilized template assays using yeast nuclear extracts from a wild-type strain (WT NE) and a temperature-sensitive *srb4-138* strain (*med17-ts* NE) as shown in (A). The native *HIS4* core promoter and upstream activating sequence were used. Recombinant activator (rGcn4) was added where applicable. 5% of the supernatant after heat inactivation and subsequent centrifugation were used as input controls (lane 1 and 2).

### 3.3.3 A label-free mass spectrometry approach to study PIC assembly

To study in more detail PIC assembly *in vitro*, we established the analysis of immobilized template eluates by tandem mass spectrometry and subsequent label-free quantitation using spectral counting. A similar method had been previously used to characterize RNA polymerases (Mosley et al, 2010) and mammalian Mediator (Paoletti et al, 2006). However, in these studies the complexes were purified by chromatography and consequently contained only a very limited number of proteins. Based on a previous mass spectrometry analysis of immobilized template eluates (Kim et al, 2007), we expected our samples to contain more than a hundred different polypeptides, including many unspecific DNA binding factors. Therefore, we first validated our label-free mass spectrometry approach by analyzing three replicates of immobilized template assays using the native HIS4 promoter template, recombinant Gcn4 and a wild-type yeast nuclear extract. 178 polypeptides were reliably detected in all three replicates, comprising 143 nuclear proteins, 23 cytosolic proteins and 12 other or unknown proteins (Figure 20). 118 proteins are annotated with the Gene ontology term “Transcription” and 85 proteins with the more specific term “Transcription from RNA polymerase II promoter”. Proteins not involved in Pol II transcription include DNA repair factors, chromosome maintenance factors as well as Pol I and Pol III components. In total 10 out of 12 Pol II subunits and 22 out of 25 Mediator subunits were reliably detected in all three replicates (Table 21). Similarly, most of the subunits of other multiprotein complexes involved in regulation of transcription initiation, like TFIID (13 out of 17 subunits), SAGA (18 out of 20 subunits), SLIK (SAGA-like complex, 15 out of 17 subunits), SWI/SNF (10 out of 12 subunits) and NuA4 (11 out of 13 subunits) were detected. No factors involved in transcription elongation, mRNA processing and termination were found.



**Figure 20: Localization of proteins identified by tandem mass spectrometry**

Proteins were classified according to their localization using the Gene ontology term finder tool (<http://go.princeton.edu>). Other includes factors that are associated with specific organelles like mitochondria or that are not annotated.

**Table 21: Mediator and Pol II subunits detected by tandem mass spectrometry analysis of wild-type immobilized template eluates**

Complex	Module	Subunit	MW	IT 1 <sup>c</sup>	IT 2 <sup>c</sup>	IT 3 <sup>c</sup>	Mean	SD
Mediator	Head	Med6	33 kDa	3	9	4	<b>5</b>	3
		Med8	25 kDa	27	22	18	<b>22</b>	4
		Med11	15 kDa	5	7	7	<b>6</b>	1
		Med17	78 kDa	4	2	9	<b>5</b>	4
		Med18	34 kDa	8	18	14	<b>13</b>	5
		Med20	23 kDa	16	29	22	<b>22</b>	6
		Med22	14 kDa	5	7	4	<b>5</b>	2
	Head/Middle	Med19	24 kDa	nd	nd	nd	-	-
	Middle	Med1	64 kDa	16	12	12	<b>13</b>	3
		Med4	32 kDa	11	16	16	<b>14</b>	3
		Med7	26 kDa	3	4	2	<b>3</b>	1
		Med9	17 kDa	10	15	17	<b>14</b>	4
		Med10	18 kDa	8	2	7	<b>6</b>	3
		Med21	16 kDa	12	11	14	<b>12</b>	2
		Med31	15 kDa	1	3	3	<b>2</b>	1
	Middle/Tail	Med14	123 kDa	21	22	32	<b>25</b>	6
	Tail	Med2/29	48 kDa	12	10	9	<b>10</b>	1
		Med3/27	43 kDa	1	3	1	<b>2</b>	1
		Med5/24	129 kDa	37	34	37	<b>36</b>	2
		Med15	120 kDa	15	12	22	<b>16</b>	5
		Med16	111 kDa	23	17	23	<b>21</b>	4
	Kinase	Med12	167 kDa	9	11	10	<b>10</b>	1
		Med13	160 kDa	4	6	3	<b>4</b>	2
Cdk8		63 kDa	nd	nd	nd	-	-	
CycC		38 kDa	1	5	4	<b>3</b>	2	
Pol II	Specific <sup>a</sup>	Rpb1	192 kDa	8	13	7	<b>9</b>	3
		Rpb2	139 kDa	7	14	12	<b>11</b>	4
		Rpb3	35 kDa	nd	nd	nd	-	-
		Rpb4	25 kDa	1	1	2	<b>1</b>	1
		Rpb7	19 kDa	2	2	4	<b>3</b>	1
		Rpb9	14 kDa	4	4	3	<b>4</b>	1
		Rpb11	14 kDa	3	5	4	<b>4</b>	1
	Shared <sup>b</sup>	Rpb5	25 kDa	18	24	19	<b>20</b>	3
		Rpb6	18 kDa	2	3	2	<b>2</b>	1
		Rpb8	17 kDa	7	8	9	<b>8</b>	1
		Rpb10	8 kDa	5	9	2	<b>5</b>	4
		Rpb12	8 kDa	3	3	3	<b>3</b>	0

<sup>a</sup> specific subunits only found in Pol II

<sup>b</sup> shared subunits between Pol I, Pol II and Pol III

<sup>c</sup> IT1-3 lists the spectral counts for the respective subunit in three replicates

<sup>d</sup> nd = subunit not detected by tandem mass spectrometry

It is important to notice that this unbiased quantitation by tandem mass spectrometry and spectral counting cannot be used for absolute quantitation. Spectral counts for a given protein depend on various factors, including the number of peptides and consequently also on the molecular weight of the protein. Small proteins are typically more difficult to detect. Therefore, not detecting a protein by tandem mass spectrometry does not necessarily mean it is absent. However, a very high linear correlation between spectral counts and relative abundance was shown (Zhu et al, 2010). Consequently, a relative quantification between different samples is possible and was used in the following experiments.

#### **3.3.4 Mediator body is recruited independent of head module**

Next, we wanted to confirm the specificity of Mediator recruitment to the promoter template. We performed immobilized template assays in a wild-type nuclear extract with and without addition of recombinant activator and analyzed the eluates by tandem mass spectrometry.

Consistent with Western blotting analysis (Figure 19B, lanes 3 and 4), Mediator subunits were only detected when Gcn4 was added (Figure 21A) confirming an activator dependent recruitment of Mediator to the promoter template. Next, we performed immobilized template assays in a *med17-ts* nuclear extract with and without addition of recombinant head module. Consistent with Western blotting analysis (Figure 19B, lanes 5 and 6), Mediator body was recruited to the promoter even in the absence of a functional head module (Figure 21A). None of the seven head subunits, but most of the middle and tail subunits were detected. At this point, the presence of the kinase module remains inconclusive due to the low number of spectral counts in all samples. Addition of recombinant head module restored Mediator integrity. Subunits from head, middle and tail were detected (Figure 21A). These results clearly demonstrate the recruitment of a stable Mediator body comprising the middle, tail and kinase module to the promoter even in the absence of functional head module.

<b>A</b>				<b>B</b>			
Module	Subunit	Act	Head	Complex	Act	Head	
Head	Med6	1.0	1.0	Gcn4	1.0	0.0	
	Med8	1.0	1.0	Rap1	-0.2	-0.1	
	Med11	1.0	1.0	Mediator head	1.0	1.0	
	Med17	1.0	1.0	Mediator body	1.0	-0.3	
	Med18	1.0	1.0	Pol II	1.0	1.0	
	Med20	1.0	1.0	TFIID	0.2	-0.3	
	Med22	1.0	1.0	SAGA	1.0	0.3	
Head/Middle	Med19	nd	nd	SWI/SNF	0.9	0.6	
Middle	Med1	1.0	-0.6	NuA4	0.9	0.1	
	Med4	1.0	-0.3	RSC	-0.5	0.2	
	Med7	1.0	-0.2	INO80	-0.2	0.1	
	Med9	nd	0.0				
	Med10	1.0	0.0				
	Med21	1.0	-0.1				
	Med31	nd	nd				
Middle/Tail	Med14	1.0	-0.4				
Tail	Med2/29	nd	0.1				
	Med3/27	1.0	-0.4				
	Med5/24	1.0	-0.5				
	Med15	1.0	-0.3				
	Med16	1.0	-0.6				
Kinase	Med12	nd	1.0				
	Med13	nd	-0.8				
	Cdk8	nd	nd				
	CycC	1.0	-0.2				

**Figure 21: Analysis of immobilized template assays by tandem mass spectrometry**

Immobilized template assays were performed in wild-type nuclear extracts with and without recombinant Gcn4; and in *med17-ts* nuclear extracts with and without recombinant head module (recombinant Gcn4 was added to both samples). Activator dependent enrichment (Act) and head dependent enrichment (Head) were calculated (1 = positive correlation; 0 = no change; 0 = negative correlation; nd = subunit was not detected in any of the samples). **(A)** spectral counts for each subunit were used for calculation. **(B)** Sum of spectral counts for all unique subunits of the respective complex were used for calculation. In addition the gene-specific activators Gcn4 and Rap1, which are directly recruited to the HIS4 UAS are shown.

### 3.3.5 Head module is only required for the recruitment of the basal machinery

Next, we asked which factors depend on the head module for promoter recruitment. Due to the lack of recombinant head module and appropriate assays, a dissection of the recruitment dependencies of various coactivator complex was not possible in the past. As expected from previous studies and Western Blot analysis (Figure 19B, lanes 3 and 4), Pol II recruitment was abolished in *med17-ts* nuclear extracts. Consistent with the previously observed rescue of *in vitro* transcription (Figure 18) and PIC formation (Figure 19B), recruitment of Pol II was restored upon addition of recombinant head module (Figure 21B). Interestingly, recruitment of many cofactors was not affected by *med17-ts* (TFIID, SAGA, NuA4, RSC and Ino80). Only SWI/SNF recruitment showed a weak dependence on Mediator head (Figure 21B).

### 3.3.6 Discussion

Taken together, our results demonstrate that Mediator is a modular coactivator complex, which can be stably recruited by gene-specific activators to a yeast promoter even in the absence of the essential head module. This observation is in contrast with previous ChIP studies of Mediator promoter occupancy in a *med17-ts* strain. At the non-permissive temperature only the tail module (Ansari et al, 2009) or no Mediator at all (Takagi et al, 2006) was observed at the promoter. The lack of add-back experiments to rule out secondary effects, the difficulty to chromatin immunoprecipitate Mediator at standard growth conditions (Fan & Struhl, 2009) and the monitoring of only one representative subunit for each module might explain these differences. Furthermore, our results demonstrate that the previously shown interdependent recruitment of SAGA, SWI/SNF and Mediator to the core-promoter (Qiu et al, 2005) does not require the head module. Our results demonstrate that despite the presence of Mediator body, TFIID, SAGA and SWI/SNF at the promoter, PIC formation does not occur. Only addition of recombinant head module restores Pol II recruitment and consequently PIC formation. This is particularly intriguing since TFIID and SAGA are also recruited by activators and were suggested to facilitate alternative PIC assemblies, which might not rely on Mediator (reviewed in (Sikorski & Buratowski, 2009)). This suggests that although promoters might differ in their dependence on TFIID and SAGA (Basehoar et al, 2004; Huisinga & Pugh, 2004) functional Mediator is always required.

This is consistent with a previously suggested role of Mediator as a general transcription factor (Takagi & Kornberg, 2006) and the global shutdown of Pol II transcription in a *med17-ts* strain (Holstege et al, 1998; Thompson & Young, 1995). Another intriguing observation is the presence of TFIID on the HIS4 promoter in the absence of Gcn4 and other cofactors (Figure 21B). Previous studies had already shown that TFIID can be directly recruited to promoters through interaction with the Rap1 transcription factor thereby facilitating PIC formation (Papai et al, 2010). Consistently, we found a predicted Rap1 binding site in the HIS4 UAS and detected Rap1 in all samples by mass spectrometry (Figure 21B). However, the fact that Rap1 recruits TFIID without facilitating PIC formation might suggest the existence of a two-step mechanism during PIC formation. At first, TFIID is loaded to the promoter. Afterwards, other gene-specific activators recruit the remaining cofactors like Mediator thereby facilitating PIC assembly and transcription initiation. Additional experiments and cross-validation of the observations have to be done to validate a sequential PIC assembly at the HIS4 promoter.

## 3.4 Additional contributions

### 3.4.1 Overview

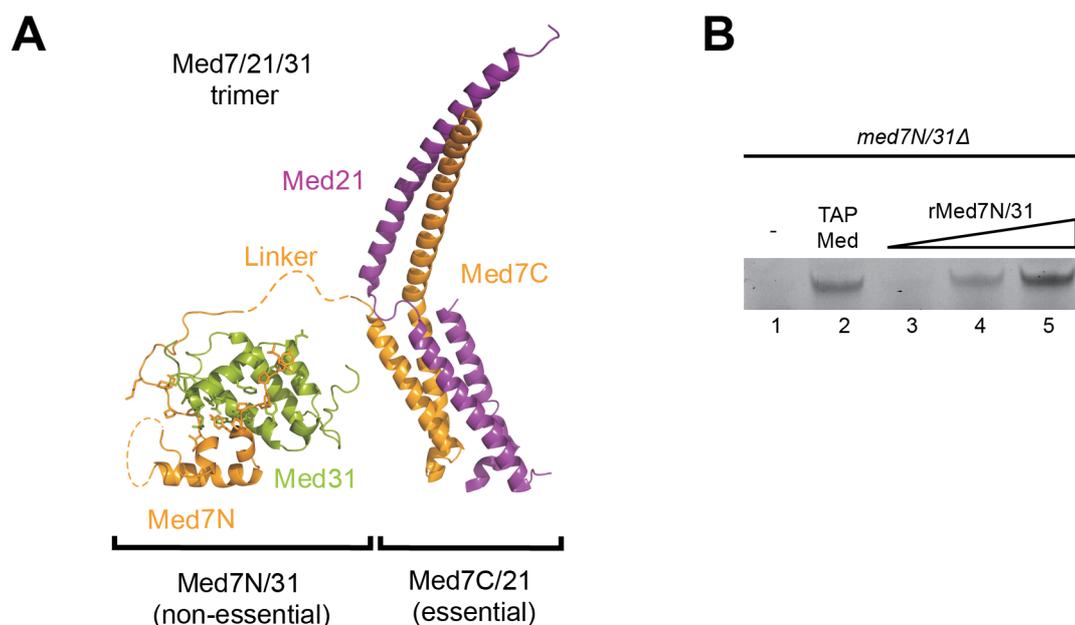
The following chapters describe additional experimental results obtained during this thesis, which lead to co-author publications. Each chapter starts with a short background overview and the relevant biological question. Afterwards, the specific experimental contributions to the respective publication are described in detail. Additional experiments are only shown and discussed when required for understanding the context. All results, detailed methods and a broad discussion of the findings can be found in the respective publication and/or thesis of the first author.

### 3.4.2 Mediator middle submodule Med7N/31 cooperates with TFIIS during activated transcription *in vitro*

Recombinant coexpression, copurification and subsequent limited proteolysis of the trimeric Med7/Med21/Med31 Mediator middle subcomplex had revealed the existence of two stable subcomplexes. On the one hand, a dimeric subcomplex comprising Med21 and the C-terminus of Med7 (Med7C; residues 103-222) which had been previously characterized (Baumli et al, 2005) and on the other hand, a dimeric subcomplex comprising Med31 and the N-terminus of Med7 (Med7N; residues 1-83) (Figure 22A). The Mediator middle subunit Med31 is particularly interesting since it is not essential for yeast viability, but highly conserved from yeast to human (Bourbon, 2008). Several large-scale studies had already shown that Med31 is required for an efficient stress response, e.g. during ethanol, cycloheximide, hydroxyurea and heat stress (Alamgir et al, 2010; Teixeira et al, 2009). Furthermore, Med31 had been shown to interact genetically with the general transcription factor TFIIS (Malagon et al, 2004) and had been implicated to act in conjunction with TFIIS during PIC formation (Guglielmi et al, 2007). In this study, we used a combination of structural biology, yeast genetics, biochemical assays and global gene expression profiling to demonstrate that Med7N/31 is a functional submodule within Mediator required for the expression of a specific subset of genes.

To characterize the function of Med7N/31 during activated transcription and its functional cooperativity with TFIIS, we performed *in vitro* transcription assays with yeast nuclear extracts. We used a plasmid-based template comprising the yeast HIS4

core promoter and a single upstream Gal4-binding site (Ranish et al, 1999). This template enabled us to use two different recombinant fusion activators. Both comprised the DNA binding domain of Gal4 and additionally either the activation domain of the strong viral activator VP16 (Gal4-VP16) or the acidic yeast activator Gal4 (Gal4-AH).



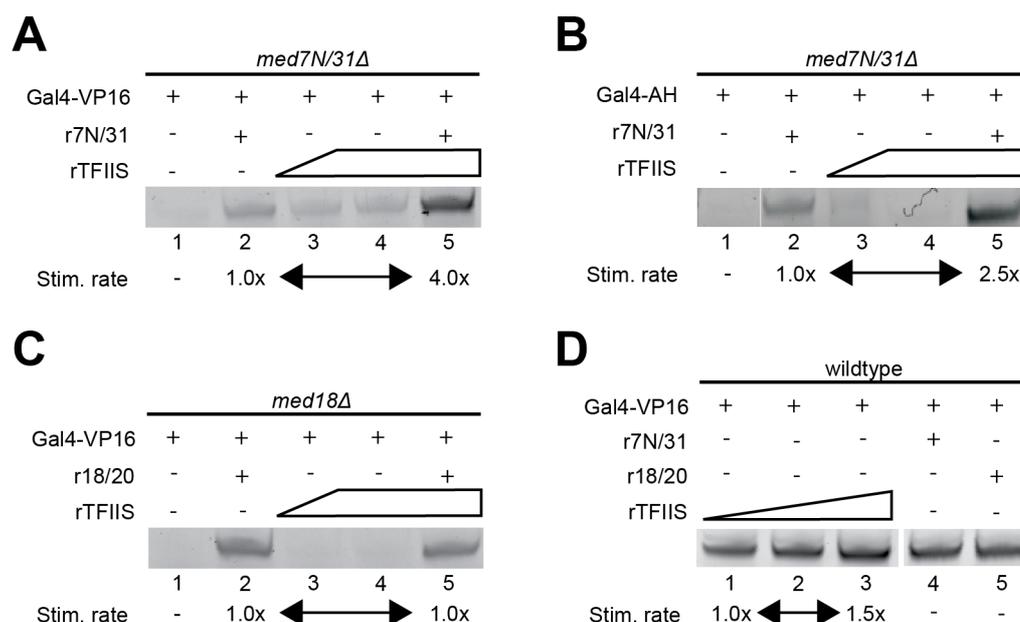
**Figure 22: Med7N/31 is a functional submodule *in vitro***

**(A)** Structural overview of the trimeric Med7/21/31 complex architecture. Structures of the non-essential Med7N/31 submodule (this study) and the previously described essential Med7C/21 subcomplex (Baumli et al, 2005) are drawn to scale. **(B)** *In vitro* transcription in a *med7N/31Δ* yeast nuclear activated with Gal4-VP16 is inactive (lane 1). Transcription can be rescued by the addition of TAP-purified Mediator (0.2 pmol, lane 2) or recombinant Med7N/31 (2–200 pmol, lanes 3–5).

First, we tested whether Med7N/31 is required for activated transcription *in vitro*. A nuclear extract from the *med7N/31Δ* strain was defective in transcription activated with either activator (Figure 22, lane 1; Figure 23B, lane 2). Addition of TAP-purified Mediator restored transcription (Figure 22, lane 2), providing a positive control. Transcription could also be restored by the addition of recombinant Med7N/31, demonstrating that the subcomplex can act in *trans* without being covalently tethered to the Mediator complex (Figure 22, lane 3-5; Figure 23B, lane 2).

Next, we tested whether Med7N/31 cooperates with TFIIIS during activated transcription *in vitro*. Indeed, transcription in a *med7N/31Δ* nuclear extract that was rescued by recombinant Med7N/31 could be enhanced by the addition of recombinant TFIIIS approximately 4x and 2.5x for Gal4-VP16 (Figure 23A) and Gal4-AH (Figure 23B) activated transcription, respectively. In wild-type nuclear extracts, TFIIIS addition

had only a weak stimulatory effect (1.5x; Figure 23D, lane 1-3) and addition of Med7N/31 had no effect (Figure 23D, lane 4). Interestingly, addition of TFIIS to a *med7N/31Δ* extract could partially rescue transcription even in the absence of recombinant Med7N/31 (Figure 23A, lane 3-4). TFIIS was not limiting in the extracts, as a three-fold higher concentration did not further increase the signal. To determine whether the observed effect is specific for the Med7N/31 submodule, we prepared a nuclear extract from the Mediator head subunit deletion strain *med18Δ*. Mediator in the *med18Δ* extract specifically lacks the Med18/20 submodule, is inactive for transcription *in vitro* and can be rescued by addition of recombinant Med18/20 in *trans* (see Chapter 3.1). In contrast to the *med7N/31Δ* extract, transcription in the *med18Δ* nuclear extract was not enhanced by TFIIS addition (Figure 23C, lane 3 and 4). Furthermore, the *med18Δ* nuclear extract, which was rescued by the addition of recombinant Med18/20, was not further stimulated by the addition of recombinant TFIIS (Figure 23C, lane 5).



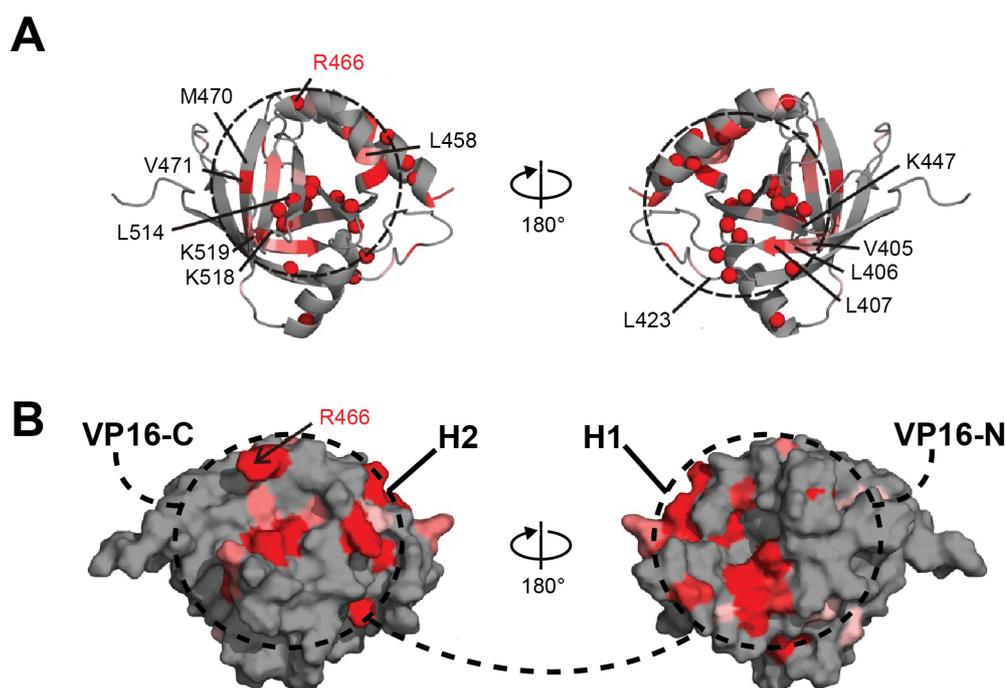
**Figure 23: Med7N/31 cooperates with TFIIS during activated transcription *in vitro***

**(A)** *In vitro* transcription in a *med7N/31Δ* yeast nuclear extract supplemented with Gal4-VP16 is inactive. Addition of recombinant Med7N/31 (r7N/31; 200 pmol; lane 2) could rescue transcription. Addition of recombinant TFIIS (rTFIIS; 20 and 60 pmol; lane 3 and 4, respectively) can partially compensate for the loss of Med7N/31. **(B)** *In vitro* transcription as described for (A) but supplemented with Gal4-AH instead of Gal4-VP16. **(C)** *In vitro* transcription in a *med18Δ* yeast nuclear extract supplemented with Gal4-VP16 is inactive. Addition of recombinant Med18/20 (r18/20; 20 pmol; lane 2) could rescue transcription. Addition of rTFIIS (20 and 60 pmol; lane 3 and 4, respectively) cannot compensate the loss of Med18/20. **(D)** *In vitro* transcription in a wild-type yeast nuclear extract supplemented with Gal4-VP16 is active (lane 1). Addition of rTFIIS has only a weak stimulatory effect (20 and 60 pmol; lane 2 and 3, respectively). Addition of r18/20 (20 pmol; lane 4) or r7N/31 (200 pmol; lane 5) to a wild-type yeast nuclear extract had no effect.

Taken together, our results demonstrate that TFIS can partially compensate for the loss of Med7N/31, and that highest transcript levels required the addition of both TFIS and Med7N/31. These observations were specific for the *med7N/31Δ* extract, as this effect was not obtained in corresponding experiments with a *med18Δ* extract. These results demonstrate the specific functional cooperativity between the Mediator submodule Med7N/31 and TFIS during Pol II transcription initiation in an *in vitro* system.

### 3.4.3 The archetypical activator VP16 targets the Mediator subunit Med25 through a conserved synergistic use of subdomains

The herpes simplex virus protein 16 (VP16) is an archetypical acidic activator, which activates expression of immediate early viral genes during infection. The C-terminal VP16 transactivation domain (TAD; residues 410-490) is widely used for studying transcription activation, typically by fusing it to the DNA-binding domain of Gal4 (Gal4-VP16). The VP16 TAD contains two functional subdomains, H1 (residues 410-452) and H2 (residues 453-490), which activate transcription independently. VP16 facilitates PIC formation through targeting several GTF, including TFIIA, TFIIB, TFIID, and TFIIH as well as Mediator. The metazoan Mediator subunit Med25 was shown to directly interact with VP16 TAD through the activator interaction domain (ACID). In this study we solved the NMR structure of Med25 ACID. Chemical shift perturbation experiments indicated that VP16 TAD interacts with an extended surface of Med25 (Figure 24A,B; for details see (Vojnić et al, 2011)).

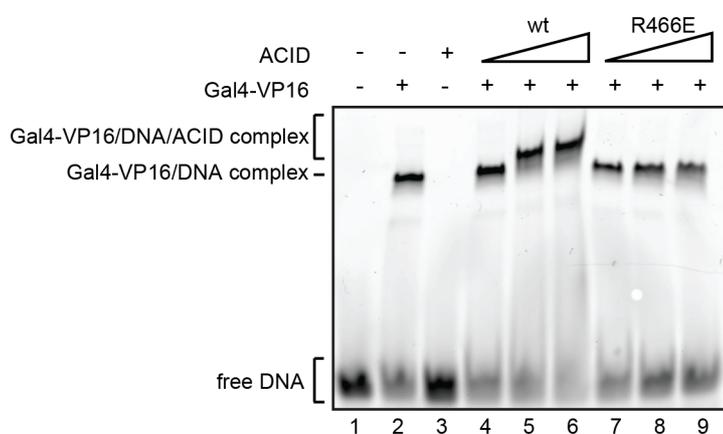


**Figure 24: NMR structure of Med25 and interaction with VP16 TAD**

(A) Ribbon model representation of the Med25 ACID NMR structure. The two views are related by a 180° rotation around the vertical axis. Residues that undergo chemical shift changes ( $\Delta\delta > 0.6$  ppm) in the  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC spectra upon VP16 TAD binding onto the ACID structure are highlighted. Rising red color intensities correspond to increasing chemical shift changes. Spheres indicate residues with signals that show binding in intermediate exchange. (B) Surface representation of the Med25 ACID structure shown in the same orientations and colors as in (A). Dashed lines indicate binding surfaces of VP16-H1 and VP16-H2 activation subdomains as determined by chemical shift perturbation experiments. Interaction "hotspot" R466, identified by biochemical assays, is highlighted.

To biochemically validate the observations from NMR chemical shift perturbation experiments and further characterize the VP16 activation mechanism, we established and performed several biochemical assays.

First, we validated the VP16-ACID interaction with electrophoretic mobility shift assays (EMSA). A fluorescently labeled DNA encompassing a Gal4-binding site enabled us to monitor formation of a binary complex of Gal4-VP16 with DNA (Figure 25, lane 2), and also formation of a ternary complex of the Gal4-VP16/DNA complex with ACID (Figure 25, lane 4-6). VP16-ACID interaction and consequently ternary complex formation could be observed as a supershift in the EMSA. To probe structural determinants in the interface, we mutated sites on ACID that showed strong chemical shift perturbations in the NMR titration experiments. All ACID variants still bound VP16 in the EMSA (data not shown), except the charge-reversal variant R466E, which completely abolished the supershift (Figure 25, lane 7-9). These results demonstrate that VP16 binding to ACID is robust and probably relies on multiple redundant contacts, and show that ACID residue R466 forms a critical interaction 'hot spot'. This basic residue lies within the H2-binding face (Figure 24B) and is conserved among ACID homologues.

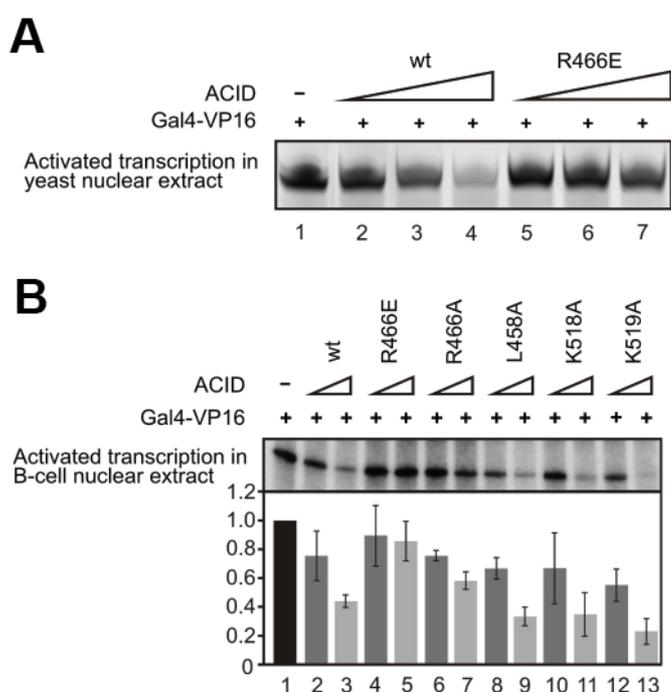


**Figure 25: ACID-VP16 interaction assay**

EMSA supershift assay reveals the importance of ACID residue R466 for VP16 binding. The binary complex formed by DNA and Gal4-VP16 (lane 2) undergoes a supershift upon addition of increasing concentrations of wild-type ACID (lanes 4-6). This supershift is abolished by ACID point mutation R466E (lane 7-9).

Next, we established a quenching assay to investigate whether the characterized VP16-ACID interaction is of functional significance during activated transcription. We performed *in vitro* activated transcription assays with yeast nuclear extracts on the

widely used HIS4 promoter with an upstream Gal4-binding site (Ranish et al, 1999). The Gal4-VP16 fusion protein is a potent transcriptional activator in this system (Figure 26A, lane 1). Addition of recombinant ACID quenched the transcription signal through an apparent competition with VP16 targets in the yeast extract (Figure 26A, lane 2-4). Thus this assay monitored interference of human ACID with functional interactions between VP16 and the basal Pol II machinery during activated transcription *in vitro*. The ACID variant R466E had the strongest effect and only quenched transcription to a low extent (Figure 26A, lane 5-7), consistent with its critical role in VP16 binding as observed in the EMSA assay. Similar experiments with a mammalian B-cell *in vitro* transcription system showed the same effects (Figure 26B; done by collaborators).

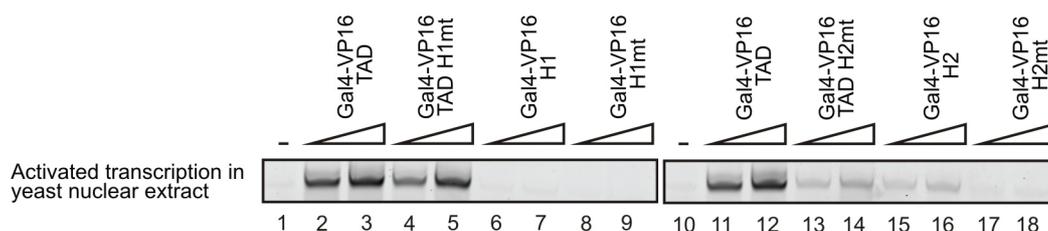


**Figure 26: VP16 activation is quenched by ACID in yeast and mammalian system**

**(A)** ACID quenches VP16 activation in a yeast *in vitro* transcription system. Assays were performed with wild-type yeast nuclear extracts (lane 1). Increasing amounts (10, 100 or 400 pmol) of either wild-type ACID (lanes 2-4) or ACID variant R466E were added (lanes 5-7). Transcription is quenched with recombinant ACID, but not with ACID variant R466E. **(B)** ACID quenches VP16 activation in a mammalian *in vitro* transcription system. Assays were performed with B-cell nuclear extracts (lane 1). Increasing amounts (530 or 850 pmol) of either wild-type ACID (lanes 2, 3) or specific ACID point mutant variants were added (lanes 4-13). ACID R466E variant hardly quenches transcription (lanes 4, 5), whereas other ACID variants do to various extents (lanes 6-13). Data are presented as average values of three experiments  $\pm$  S.D., and one representative gel is shown.

To evaluate the contributions of the VP16 subdomains to transcription activation in yeast, we prepared fusion proteins of the Gal4 DNA-binding domain with H1, H2, mutated H1 (H1mt) and mutated H2 (H2mt). The mutations targeted functionally required hydrophobic residues. Additionally fusions with TAD that was mutated at functionally required hydrophobic residues in H1 or H2 (TAD H1mt and TAD H2mt, respectively) were prepared. In these assays, H2, but not H1, was alone able to activate transcription, although weakly (Figure 27, lanes 6, 7, 15, 16). Consistently,

TAD that carried a mutation in H2 supported activated transcription only weakly (Figure 27, lanes 13, 14), whereas TAD that carried a mutation in H1 strongly activated transcription, to nearly the levels of wild-type TAD (Figure 27, lanes 4 and 5). Thus, the VP16 TAD subdomains H1 and H2 cooperate during activated transcription in the yeast system, with the main contribution coming from H2.



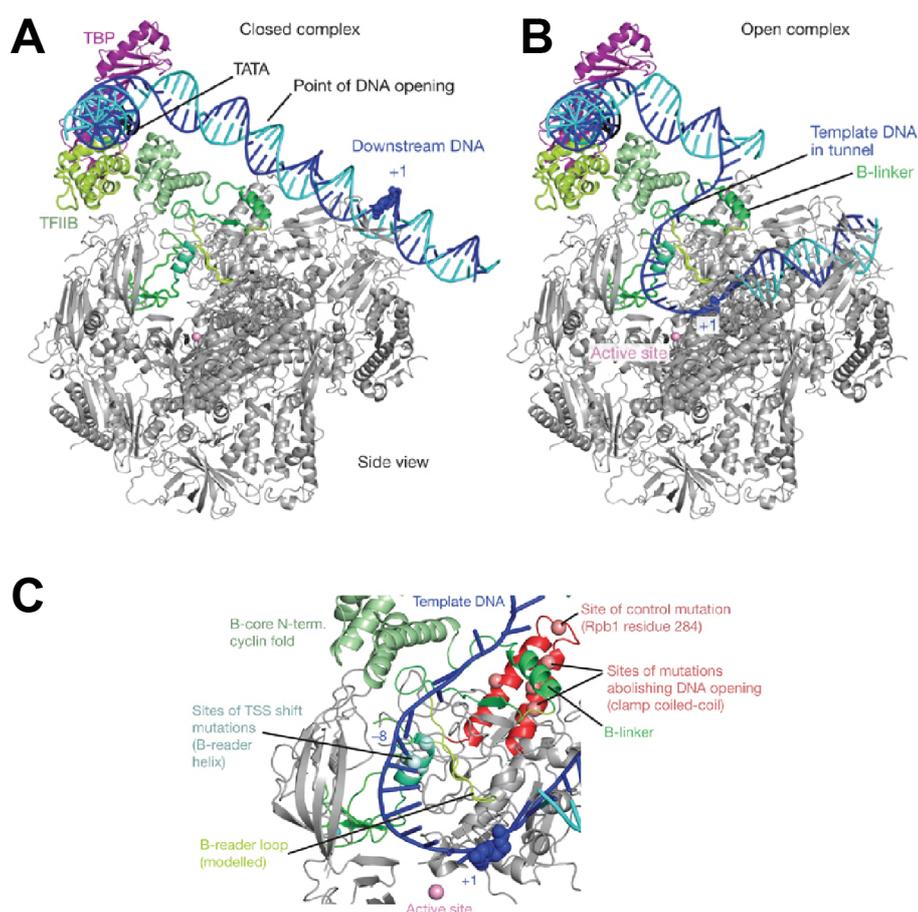
**Figure 27: VP16 subdomains H1 and H2 activate yeast transcription synergistically**

In vitro transcription with yeast nuclear extracts Transcription in yeast nuclear extracts was monitored in the absence (lanes 1, 10) or the presence of different Gal4-VP16 variants, including Gal4 fusions with VP16 TAD (lanes 2, 3, 11, 12), TAD carrying the H1 subdomain mutation F442P (H1mt, lanes 4, 5), H1 (lanes 6, 7), H1mt (lanes 8, 9), TAD carrying the H2 subdomain mutations F473A, F475A, and F479A (H2mt, lanes 13, 14), H2 (lanes 15, 16), and H2mt (lanes 17, 18).

Taken together, our results characterize the interaction of the metazoan Mediator subunit Med25 with the archetypical acidic transcription factor VP16. Strikingly, the synergistic activation mode of VP16 is conserved, although it binds different target proteins in the yeast and human systems. Yeast Mediator contains neither Med25 nor another subunit with ACID homology. In both systems VP16 TAD uses its two subdomains H1 and H2 in a synergistic manner, indicating that the yeast transcription machinery contains at least two non-overlapping sites capable of binding to the two TAD subdomains. Our results contribute to the understanding of how activation domains may have evolved to adapt to different unrelated target surfaces.

### 3.4.4 General transcription factor TFIIIB controls transcription start site selection

The general transcription factor TFIIIB plays an important role during Pol II transcription initiation. TFIIIB recruits Pol II to the promoter with its N-terminal zinc-ribbon domain (B-ribbon) (Bushnell et al, 2004; Chen & Hahn, 2003), and at same time contacts TBP and DNA through its C-terminal domain (B-core) (Nikolov et al, 1995).



**Figure 28: Models of the Pol II closed and open complex**

**(A)** Model of the closed complex (minimal PIC) based on the Pol II-TFIIIB crystal structure. DNA template and non-template strands are in blue and cyan, respectively. The TATA element is in black and the nucleotide in the template strand that represents position 11 in the open complex is shown as a space-filling model. **(B)** Model of the open complex (minimal PIC) in the same orientation as in (A). **(C)** Location of nucleotides in DNA template strand initiator consensus sequence and mutations influencing start site selection and DNA opening. The open complex model is shown around the active centre. Positions -8 and +1 of the template strand are labelled. Position +8 lies adjacent to the B-reader helix that contains residues important for TSS selection (Glu 62, Trp 63, Arg 64, Phe 66, pale green spheres). The mobile B-reader loop (green-yellow), which contains residues Arg 78 and Val 79 required for initial transcription and TSS selection, could reach near positions -1 and +1. Sites of mutations abolishing DNA opening in archaeal transcription are shown as salmon spheres.

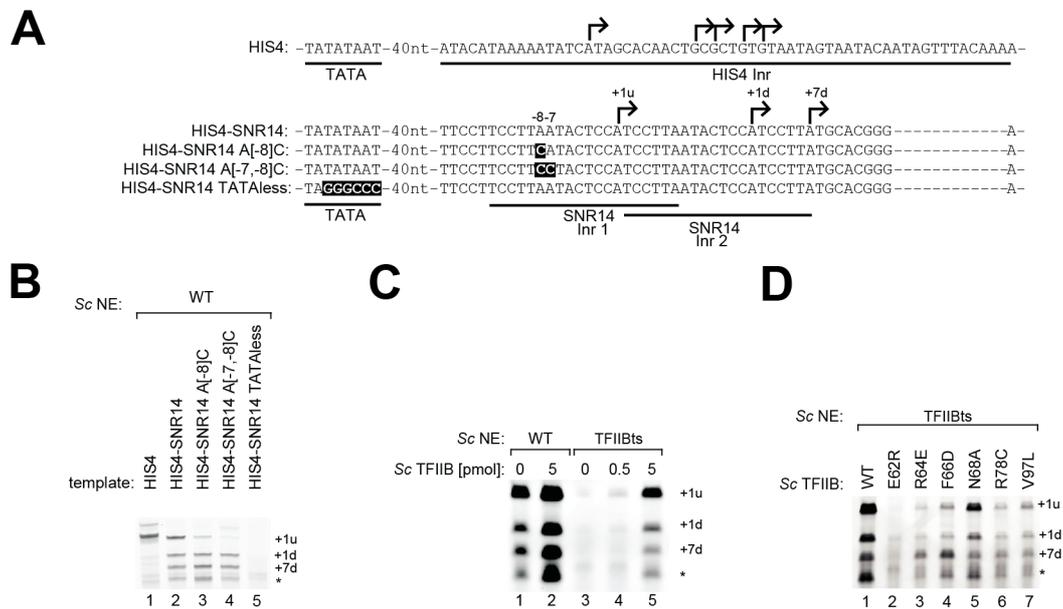
Interestingly, TFIIB has an additional postrecruitment function after PIC formation. Mutations in the region connecting the B-ribbon and the B-core had been previously shown to affect transcription and transcription start site (TSS) selection, but not PIC formation. Several lines of evidence indicated that yeast Pol II scans the DNA for an INR sequence motif that defines the TSS. This scanning apparently involves threading of the DNA template strand through the template tunnel until an INR is detected by sequence-specific interactions. Since the template tunnel passes the active site, INR detection must occur near the active site where RNA synthesis is initiated. Previous studies had shown that TSS shifts induced by TFIIB mutations depend on the INR sequence and that Pol II and TFIIB are alone responsible for TSS selection.

In this study we solved the crystal structure of a Pol II-TFIIB complex at 4.3 Å and proposed a model for the Pol II closed and open complex (Figure 28A and B, respectively). Furthermore, two novel structural elements in TFIIB were identified, the 'B-linker' and the 'B-reader' element (Figure 28C). *In vitro* biochemical assays demonstrated that the B-linker is involved in promoter opening (not shown). Strikingly, the highly conserved B-reader flanks the template tunnel in our open complex model (Figure 28B and C), and mutations reported to specifically affect TSS selection mapped to this region (Figure 28C). Therefore we hypothesized that the B-reader is involved in INR recognition and TSS selection during Pol II scanning.

To test the scanning model and to demonstrate a B-reader dependent INR recognition mechanism, we designed a specialized *in vitro* transcription template. This template comprised the yeast HIS4 core promoter sequence around the TATA box, a single upstream Gal4-DNA binding site and a duplicated INR of the yeast SNR14 (Figure 29A). This fusion template generated transcripts that initiated at the previously mapped TSS of the SNR14 promoter (Figure 29B, lane 2). Thus, the INR alone determined the TSS, consistent with the scanning model. The yeast INR consensus motif had been shown to comprise a conserved residue at position -8, and a CA or TG dinucleotide at positions -1/+1 of the non-template strand. Consistently, mutation of the -8 position in the first INR of the fusion promoter led to an almost complete loss of the corresponding transcript (Figure 29B, lane 3). When both positions -8 and -7 were mutated, transcription from the first TSS was totally abolished (Figure 29B, lane 4). Mutation of the TATA box abolished transcription from all TSS (Figure 29B, lane 5). In the open complex model, the complementary template T residue at -8 is adjacent to the B-reader (Figure 28B and C). To test mutations in TFIIB, we prepared a nuclear extract from a

strain carrying a temperature-sensitive mutation in TFIIB (*TFIIB-ts*). This extract is inactive in *in vitro* transcription but can be restored by addition of recombinant TFIIB (Figure 29C). We used this system to test several mutant variants of TFIIB in their ability to rescue *in vitro* transcription and changes in TSS selection. Consistent with the modelling and published data, mutation of the respective B-reader residues decreased transcription efficiency and in some cases led to detectable changes in TSS selection (Figure 29D).

Taken together, our results demonstrate that the INR alone determines TSS selection and that the -8 position is important for the INR recognition. Furthermore, our results suggest that the INR is recognized with the help of the B-reader during Pol II scanning.



**Figure 29: B-reader and transcription start site selection**

(A) Design of the fusion promoter template HIS4-SNR14. In promoters A[-8]C and A[-7, -8]C, adenines at position -8 or at -8 and -7, respectively, relative to the first TSS were replaced by cytosines (white). Arrows indicate *in vitro* TSSs in HIS4 and HIS4-SNR14 fusion promoters, respectively. (B) Three different TSSs are used during *in vitro* transcription in yeast extracts with the HIS4-SNR14 fusion promoter (lane 2). The mutations in promoters A[-8]C and A[-7,-8]C gradually eliminate recognition of the first TSS (lanes 3 and 4). (C) Nuclear extract from a temperature-sensitive yeast strain with a mutation in TFIIB (TFIIBts) is essentially inactive in *in vitro* transcription (lane 3). Activity is restored when recombinant wild-type (WT) TFIIB is added (lane 5). Asterisk denotes a non-specific band. The HIS4-SNR14 template and the location of the observed TSSs (+1 upstream (u), +1 downstream (d), +7 downstream) (D) Mutations in the yeast B-reader residues Arg 64, Phe 66, Arg 78 and Val 79, but not Asn 68, lead to TSS shifts.

## 4 Conclusion and Outlook

The head module of the general coactivator complex Mediator plays an important role during Pol II transcription initiation. In this study, we used a structure-function-system correlation, combining X-ray crystallography, yeast genetics, biochemical assays, chromatin immunoprecipitation, genome-wide expression profiling and label-free mass spectrometry, to further characterize its architecture and function. Small-structure guided system perturbations were used to dissect the distinct roles of the Med8C/18/20 and Med11/22 submodules without affecting head module integrity *per se*. This approach is in general superior to classic gene deletion studies since it does not affect complex integrity. Therefore, effects can be directly attributed to the mutated subunit. A similar approach was established for studying recombinant 7-subunit head modules. In the future, the combination of functional *in vitro* assays supplemented with recombinant subcomplexes and an unbiased analysis of promoter-associated factors by label-free tandem mass spectrometry will enable a precise dissection of Mediator function.

Future goals include the following:

1) *Further characterization of Med11/22 function during Pol II transcription*

In this study, structure-guided Med11/22 mutations were shown to affect the expression of at least 400 genes. In contrast, previous studies had shown that the *med17-ts* mutation in the head module affects the expression of most if not all genes. To test whether Med11/22 is also globally required for gene expression, the structure-based temperature-sensitive mutation *med11-L73K/K80E* could be used. Global synthesis rates of *med17-ts* and *med11-L73K/K80E* could be determined and compared by dynamic transcriptome analysis (Miller et al, 2011). Directly after heat inactivation the nascent transcripts would be labeled and subsequently analyzed. This approach minimizes secondary effects, since the mutation has no effect on gene expression at the permissive temperature.

Furthermore, site-directing crosslinking approaches could be used to characterize the interaction surface of Med11/22. Site-directed crosslinkers could be easily introduced into the recombinant head module. *In vitro* PIC assembly with a *med17-ts* nuclear extract and the modified recombinant head

module followed by crosslinking could potentially confirm the reported yeast-two-hybrid interaction with TFIIH and identify additional interaction partners.

2) *Functional characterization of other head module subunits*

The versatile *in vitro* assays based on *med17-ts* nuclear extracts and recombinant Mediator head module can be used in the future to characterize the remaining essential subunits, Med6, Med8 and Med17. Mutant variants of the recombinant head module could be generated and tested for the rescue of *in vitro* transcription similar to the approach used to characterize TFIIB (see Chapter 3.4.4). Furthermore PIC assembly of various mutants could be tested. Similar to the approach described above, site-directed crosslinkers could be introduced to identify interaction partners in Pol II and the GTF.

3) *Structural and functional studies of recombinant core Mediator*

Recombinant coexpression and copurification strategies for Mediator head and middle modules are already available in the laboratory. In the future, a similar strategy could be used to obtain recombinant core Mediator comprising all essential subunits. This complex could be used for high-resolution electron microscopy and crosslinking mass spectrometry. Fusion of recombinant core Mediator with a DNA-binding domain, like the bacterial *lexA*-DNA binding domain, could be used to study the role of Mediator during PIC formation isolated from other coactivator complexes.

4) *Identification and characterization of inter-module interactions within Mediator*

Although the architecture of individual Mediator modules has been characterized in detail (Koschubs et al, 2010; Takagi et al, 2006), the inter-module interactions remain largely unclear. In this study, we demonstrated the existence of a stable head-less Mediator (termed Mediator body). A recombinant head module can associate with Mediator body and thus rescue Mediator function. In the future, mutant variants of the recombinant head module could be used to identify the molecular interaction between Mediator head and body.

5) *Characterization of the interplay between Mediator and other coactivators at different Pol II promoter classes*

Recent genome-wide studies had suggested that different Pol II promoters are dominated by different coactivator complexes (Basehoar et al, 2004; Huisinga & Pugh, 2004). While TATA box containing promoters rely on SAGA, TATA-less promoters rely on TFIID for activation. In this study, we demonstrated that several coactivator complexes including Mediator, SAGA and TFIID, are recruited to a TATA box containing yeast promoter *in vitro*. In the future, a TATA-less promoter template for immobilized template assays could be established. Differences in factor recruitment as well as the role of Mediator could be determined using the unbiased label-free mass spectrometry approach established in this study.

## References

- Alamgir M, Erukova V, Jessulat M, Azizi A, Golshani A (2010) Chemical-genetic profile analysis of five inhibitory compounds in yeast. *BMC Chem Biol* **10**: 6
- Allison LA, Moyle M, Shales M, Ingles CJ (1985) Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. *Cell* **42**: 599-610
- Ansari S, He Q, Morse R (2009) Mediator complex association with constitutively transcribed genes in yeast. *Proc Natl Acad Sci U S A* **106**: 16734-16739
- Aparicio O, Geisberg JV, Sekinger E, Yang A, Moqtaderi Z, 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
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**: 25-29
- Asturias F, Jiang Y, Myers L, Gustafsson C, Kornberg R (1999) Conserved structures of mediator and RNA polymerase II holoenzyme. *Science* **283**: 985-987
- Auble DT (2009) The dynamic personality of TATA-binding protein. *Trends Biochem Sci* **34**: 49-52
- Bäckström S, Elfving N, Nilsson R, Wingsle G, Björklund S (2007) Purification of a Plant Mediator from *Arabidopsis thaliana* Identifies PFT1 as the Med25 Subunit. *Mol Cell* **26**: 717-729
- Basehoar AD, Zanton SJ, Pugh BF (2004) Identification and distinct regulation of yeast TATA box-containing genes. *Cell* **116**: 699-709
- Baumli S, Hoepfner S, Cramer P (2005) A conserved mediator hinge revealed in the structure of the MED7.MED21 (Med7.Srb7) heterodimer. *J Biol Chem* **280**: 18171-18178
- Becker PB, Hörz W (2002) ATP-dependent nucleosome remodeling. *Annu Rev Biochem* **71**: 247-273
- Beve J, Hu G, Myers L, Balciunas D, Werngren O, Hultenby K, Wibom R, Ronne H, Gustafsson C (2005) The structural and functional role of Med5 in the yeast Mediator tail module. *J Biol Chem* **280**: 41366-41372
- Bhaumik SR (2011) Distinct regulatory mechanisms of eukaryotic transcriptional activation by SAGA and TFIID. *Biochim Biophys Acta* **1809**: 97-108
- Bhaumik SR, Raha T, Aiello DP, Green MR (2004) In vivo target of a transcriptional activator revealed by fluorescence resonance energy transfer. *Genes Dev* **18**: 333-343
- Björklund S, Gustafsson C (2005) The yeast Mediator complex and its regulation. *Trends Biochem Sci* **30**: 240-244
- Bontems F, Verger A, Dewitte F, Lens Z, Baert J-L, Ferreira E, de Launoit Y, Sizun C, Guittet E, Villeret V, Monté D (2010) NMR structure of the human Mediator MED25 ACID domain. *Journal of structural biology*
- Boube M, Joulia L, Cribbs DL, Bourbon H-M (2002) Evidence for a mediator of RNA polymerase II transcriptional regulation conserved from yeast to man. *Cell* **110**: 143-151
- Bourbon H-M (2008) Comparative genomics supports a deep evolutionary origin for the large, four-module transcriptional mediator complex. *Nucleic Acids Res*
- Boyle EI, Weng S, Gollub J, Jin H, Botstein D, Cherry JM, Sherlock G (2004) GO::TermFinder--open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics* **20**: 3710-3715

- Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248-254
- Budisa N, Steipe B, Demange P, Eckerskorn C, Kellermann J, 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 (2005) Connections between mRNA 3' end processing and transcription termination. *Curr Opin Cell Biol* **17**: 257-261
- Buratowski S (2009) Progression through the RNA polymerase II CTD cycle. *Mol Cell* **36**: 541-546
- Bushnell DA, Westover KD, Davis RE, Kornberg RD (2004) Structural basis of transcription: an RNA polymerase II-TFIIB cocystal at 4.5 Angstroms. *Science* **303**: 983-988
- Cadena DL, Dahmus ME (1987) Messenger RNA synthesis in mammalian cells is catalyzed by the phosphorylated form of RNA polymerase II. *J Biol Chem* **262**: 12468-12474
- Cai G, Imasaki T, Takagi Y, Asturias FJ (2009) Mediator structural conservation and implications for the regulation mechanism. *Structure* **17**: 559-567
- Cai G, Imasaki T, Yamada K, Cardelli F, Takagi Y, Asturias FJ (2010) Mediator head module structure and functional interactions. *Nat Struct Mol Biol* **17**: 273-279
- Cairns BR (2009) The logic of chromatin architecture and remodelling at promoters. *Nature* **461**: 193-198
- Cantin GT, Stevens JL, Berk AJ (2003) Activation domain-mediator interactions promote transcription preinitiation complex assembly on promoter DNA. *Proc Natl Acad Sci USA* **100**: 12003-12008
- Chalkley GE, Verrijzer CP (1999) DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator. *EMBO J* **18**: 4835-4845
- Chen H-T, Hahn S (2003) Binding of TFIIB to RNA polymerase II: Mapping the binding site for the TFIIB zinc ribbon domain within the preinitiation complex. *Mol Cell* **12**: 437-447
- Chen H-T, Warfield L, Hahn S (2007) The positions of TFIIF and TFIIE in the RNA polymerase II transcription preinitiation complex. *Nat Struct Mol Biol* **14**: 696-703
- Cho EJ, Buratowski S (1999) Evidence that transcription factor IIB is required for a post-assembly step in transcription initiation. *J Biol Chem* **274**: 25807-25813
- Cler E, Papai G, Schultz P, Davidson I (2009) Recent advances in understanding the structure and function of general transcription factor TFIID. *Cell Mol Life Sci* **66**: 2123-2134
- Corden JL, Cadena DL, Ahearn JM, Dahmus ME (1985) A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. *Proc Natl Acad Sci USA* **82**: 7934-7938
- Cosma MP, Panizza S, Nasmyth K (2001) Cdk1 triggers association of RNA polymerase to cell cycle promoters only after recruitment of the mediator by SBF. *Mol Cell* **7**: 1213-1220
- Cramer P (2002a) Common structural features of nucleic acid polymerases. *Bioessays* **24**: 724-729
- Cramer P (2002b) Multisubunit RNA polymerases. *Curr Opin Struct Biol* **12**: 89-97
- Crick F (1970) Central dogma of molecular biology. *Nature* **227**: 561-563
- Cuff JA, Barton GJ (1999) Evaluation and improvement of multiple sequence methods for protein secondary structure prediction. *Proteins* **34**: 508-519
- Davis J, Takagi Y, Kornberg R, Asturias F (2002) Structure of the yeast RNA polymerase II holoenzyme: Mediator conformation and polymerase interaction. *Mol Cell* **10**: 409-415

- DeLano W. (2002) The PyMOL Molecular Graphics System. *DeLano Scientific: San Carlos, CA, USA*.
- Deng W, Roberts SGE (2006) Core promoter elements recognized by transcription factor IIB. *Biochem Soc Trans* **34**: 1051-1053
- Dotson M, Yuan C, Roeder R, Myers L, Gustafsson C, Jiang Y, Li Y, Kornberg R, Asturias F (2000) Structural organization of yeast and mammalian mediator complexes. *Proc Natl Acad Sci U S A* **97**: 14307-14310
- Edgar R (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792-1797
- Elmlund H, Baraznenok V, Lindahl M, Samuelsen C, Koeck P, Holmberg S, Hebert H, Gustafsson C (2006) The cyclin-dependent kinase 8 module sterically blocks Mediator interactions with RNA polymerase II. *Proc Natl Acad Sci U S A* **103**: 15788-15793
- Emsley P, Cowtan K (2004) Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallographica Section D - Biological Crystallography* **60**
- Enyenihi AH, Saunders WS (2003) Large-scale functional genomic analysis of sporulation and meiosis in *Saccharomyces cerevisiae*. *Genetics* **163**: 47-54
- Esnault C, Ghavi-Helm Y, Brun S, Soutourina J, Van Berkum N, Boschiero C, Holstege F, Werner M (2008) Mediator-dependent recruitment of TFIID modules in preinitiation complex. *Mol Cell* **31**: 337-346
- Fan X, Struhl K (2009) Where does mediator bind in vivo? *PLoS ONE* **4**: e5029
- Field Y, Kaplan N, Fondufe-Mittendorf Y, Moore IK, Sharon E, Lubling Y, Widom J, Segal E (2008) Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. *PLoS Comput Biol* **4**: e1000216
- Firestein R, Bass A, Kim S, Dunn I, Silver S, Guney I, Freed E, Ligon A, Vena N, Ogino S, Chheda M, Tamayo P, Finn S, Shrestha Y, Boehm J, Jain S, Bojarski E, Mermel C, Barretina J, Chan J, Baselga J, Tabernero J, Root D, Fuchs C, Loda M, Shivdasani R, Meyerson M, Hahn W (2008) CDK8 is a colorectal cancer oncogene that regulates beta-catenin activity. *Nature* **455**: 547-551
- Flanagan P, Kelleher R, Sayre M, Tschochner H, Kornberg R (1991) A mediator required for activation of RNA polymerase II transcription in vitro. *Nature* **350**: 436-438
- Gadbois EL, Chao DM, Reese JC, Green MR, Young RA (1997) Functional antagonism between RNA polymerase II holoenzyme and global negative regulator NC2 in vivo. *Proc Natl Acad Sci USA* **94**: 3145-3150
- Gavin A-C, Aloy P, Grandi P, Krause R, Boesche M, Marzioch M, Rau C, Jensen LJ, Bastuck S, Dümpelfeld B, Edelmann A, Heurtier M-A, Hoffman V, Hoefert C, Klein K, Hudak M, Michon A-M, Schelder M, Schirle M, Remor M, Rudi T, Hooper S, Bauer A, Bouwmeester T, Casari G, Drewes G, Neubauer G, Rick JM, Kuster B, Bork P, Russell RB, Superti-Furga G (2006) Proteome survey reveals modularity of the yeast cell machinery. *Nature* **440**: 631-636
- Gavin A-C, Bösch M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon A-M, Cruciat C-M, Remor M, Höfert C, Schelder M, Brajenovic M, Ruffner H, Merino A, Klein K, Hudak M, Dickson D, Rudi T, Gnau V, Bauch A, Bastuck S, Huhse B, Leutwein C, Heurtier M-A, Copley RR, Edelmann A, Querfurth E, Rybin V, Drewes G, Raida M, Bouwmeester T, Bork P, Seraphin B, Kuster B, Neubauer G, Superti-Furga G (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**: 141-147
- Geiduschek EP, Ouhammouch M (2005) Archaeal transcription and its regulators. *Mol Microbiol* **56**: 1397-1407
- Geiger SR, Lorenzen K, Schrieck A, Hanecker P, Kostrewa D, Heck AJR, Cramer P (2010) RNA polymerase I contains a TFIIF-related DNA-binding subcomplex. *Mol Cell* **39**: 583-594

- Gentleman R, Carey V, Bates D, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini A, Sawitzki G, Smith C, Smyth G, Tierney L, Yang J, Zhang J (2004) Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biol* **5**: R80
- Gill G, Ptashne M (1988) Negative effect of the transcriptional activator GAL4. *Nature* **334**: 721-724
- Gouet P, Courcelle E, Stuart D, Metoz F (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**: 305-308
- Guglielmi B, Soutourina J, Esnault C, Werner M (2007) TFIIIS elongation factor and Mediator act in conjunction during transcription initiation in vivo. *Proc Natl Acad Sci U S A* **104**: 16062-16067
- Guglielmi B, van Berkum N, Klapholz B, Bijma T, Boube M, Boschiero C, Bourbon H, Holstege F, Werner M (2004) A high resolution protein interaction map of the yeast Mediator complex. *Nucleic Acids Res* **32**: 5379-5391
- Hahn S (2004) Structure and mechanism of the RNA polymerase II transcription machinery. *Nat Struct Mol Biol* **11**: 394-403
- Han S, Lee Y, Gim B, Ryu G, Park S, Lane W, Kim Y (1999) Activator-specific requirement of yeast mediator proteins for RNA polymerase II transcriptional activation. *Mol Cell Biol* **19**: 979-988
- Henry KW, Wyce A, Lo W-S, Duggan LJ, Emre NCT, Kao C-F, Pillus L, Shilatifard A, Osley MA, Berger SL (2003) Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. *Genes Dev* **17**: 2648-2663
- Herbig E, Warfield L, Fish L, Fishburn J, Knutson B, Moorefield B, Pacheco D, Hahn S (2010) Mechanism of Mediator Recruitment by Tandem Gcn4 Activation Domains and Three Gal11 Activator-Binding Domains. *Molecular and Cellular Biology* **30**: 2376
- Higuchi R, Krummel B, Saiki RK (1988) A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic acids research* **16**: 7351-7367
- Hirata A, Murakami KS (2009) Archaeal RNA polymerase. *Curr Opin Struct Biol* **19**: 724-731
- Hoepfner S, Baumli S, Cramer P (2005) Structure of the mediator subunit cyclin C and its implications for CDK8 function. *J Mol Biol* **350**: 833-842
- Holstege F, Jennings E, Wyrick J, Lee T, Hengartner C, Green M, Golub T, Lander E, Young R (1998) Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95**: 717-728
- Huang R, MAHESHWARI N, BONNER J (1960) Enzymatic synthesis of RNA. *Biochem Biophys Res Commun* **3**: 689-694
- Huisinga KL, Pugh BF (2004) A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in *Saccharomyces cerevisiae*. *Mol Cell* **13**: 573-585
- Hurwitz J, Bresler A, Diring R (1960) The enzymic incorporation of ribonucleotides into polyribonucleotides and the effect of DNA. *Biochem Biophys Res Commun* **3**: 15-18
- Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* **292**: 195-202
- Juven-Gershon T, Hsu J-Y, Theisen JW, Kadonaga JT (2008) The RNA polymerase II core promoter - the gateway to transcription. *Curr Opin Cell Biol* **20**: 253-259
- Kabsch W (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *Journal of Applied Crystallography* **26**: 795-800
- Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, Berkum NLv, Ebmeier CC, Goossens J, Rahl PB, Levine SS, Taatjes DJ, Dekker J, Young RA (2010) Mediator and cohesin connect gene expression and chromatin architecture. *Nature*: 1-6

- Kang J, Kim S, Hwang M, Han S, Lee Y, Kim Y (2001) The structural and functional organization of the yeast mediator complex. *J Biol Chem* **276**: 42003-42010
- Kaufmann J, Smale ST (1994) Direct recognition of initiator elements by a component of the transcription factor IID complex. *Genes Dev* **8**: 821-829
- Kelleher R, Flanagan P, Kornberg R (1990) A novel mediator between activator proteins and the RNA polymerase II transcription apparatus. *Cell* **61**: 1209-1215
- Kim B, Nesvizhskii A, Rani P, Hahn S, Aebersold R, Ranish J (2007) The transcription elongation factor TFIIIS is a component of RNA polymerase II preinitiation complexes. *Proc Natl Acad Sci U S A* **104**: 16068-16073
- Kim JB, Sharp PA (2001) Positive transcription elongation factor B phosphorylates hSPT5 and RNA polymerase II carboxyl-terminal domain independently of cyclin-dependent kinase-activating kinase. *J Biol Chem* **276**: 12317-12323
- Kim JL, Nikolov DB, Burley SK (1993a) Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* **365**: 520-527
- Kim TK, Ebright RH, Reinberg D (2000) Mechanism of ATP-dependent promoter melting by transcription factor IIH. *Science* **288**: 1418-1422
- Kim Y, Geiger JH, Hahn S, Sigler PB (1993b) Crystal structure of a yeast TBP/TATA-box complex. *Nature* **365**: 512-520
- Kim YJ, Björklund S, Li Y, Sayre MH, Kornberg RD (1994) A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* **77**: 599-608
- Knop M, Finger A, Braun T, Hellmuth K, Wolf DH (1996) Der1, a novel protein specifically required for endoplasmic reticulum degradation in yeast. *The EMBO journal* **15**: 753-763
- Kornberg R (2005) Mediator and the mechanism of transcriptional activation. *Trends Biochem Sci* **30**: 235-239
- Koschubs T, Lorenzen K, Baumli S, Sandström S, Heck AJR, Cramer P (2010) Preparation and topology of the Mediator middle module. *Nucleic acids research* **38**: 3186-3195
- Koschubs T, Seizl M, Larivière L, Kurth F, Baumli S, Martin DE, Cramer P (2009) Identification, structure, and functional requirement of the Mediator submodule Med7N/31. *EMBO J* **28**: 69-80
- Kostrewa D, Zeller ME, Armache K-J, Seizl M, Leike K, Thomm M, Cramer P (2009) RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature* **462**: 323-330
- Krogan NJ, Cagney G, Yu H, Zhong G, Guo X, Ignatchenko A, Li J, Pu S, Datta N, Tikuisis AP, Punna T, Peregrín-Alvarez JM, Shales M, Zhang X, Davey M, Robinson MD, Paccanaro A, Bray JE, Sheung A, Beattie B, Richards DP, Canadien V, Lalev A, Mena F, Wong P, Starostine A, Canete MM, Vlasblom J, Wu S, Orsi C, Collins SR, Chandran S, Haw R, Rilstone JJ, Gandi K, Thompson NJ, Musso G, St Onge P, Ghanny S, Lam MHY, Butland G, Altaf-Ul AM, Kanaya S, Shilatifard A, O'Shea E, Weissman JS, Ingles CJ, Hughes TR, Parkinson J, Gerstein M, Wodak SJ, Emili A, Greenblatt JF (2006) Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* **440**: 637-643
- Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680-685
- Lahmy S, Bies-Etheve N, Lagrange T (2010) Plant-specific multisubunit RNA polymerase in gene silencing. *Epigenetics* **5**: 4-8
- Larivière L, Geiger S, Hoepfner S, Röther S, Strässer K, Cramer P (2006) Structure and TBP binding of the Mediator head subcomplex Med8-Med18-Med20. *Nat Struct Mol Biol* **13**: 895-901

- Larivière L, Seizl M, van Wageningen S, Röther S, van de Pasch L, Feldmann H, Strässer K, Hahn S, Holstege FCP, Cramer P (2008) Structure-system correlation identifies a gene regulatory Mediator submodule. *Genes Dev* **22**: 872-877
- Lee T, Wyrick J, Koh S, Jennings E, Gadbois E, Young R (1998) Interplay of positive and negative regulators in transcription initiation by RNA polymerase II holoenzyme. *Mol Cell Biol* **18**: 4455-4462
- Lee TI, Causton HC, Holstege FC, Shen WC, Hannett N, Jennings EG, Winston F, Green MR, Young RA (2000) Redundant roles for the TFIID and SAGA complexes in global transcription. *Nature* **405**: 701-704
- Leurent C, Sanders S, Ruhlmann C, Malouh V, Weil PA, Kirschner DB, Tora L, Schultz P (2002) Mapping histone fold TAFs within yeast TFIID. *The EMBO journal* **21**: 3424-3433
- Li B, Carey M, Workman JL (2007) The role of chromatin during transcription. *Cell* **128**: 707-719
- Liston D, Johnson P (1999) Analysis of a Ubiquitous Promoter Element in a Primitive Eukaryote: Early Evolution of the Initiator Element. *Molecular and Cellular Biology* **19**: 2380
- Liu Y, Kung C, Fishburn J, Ansari A, Shokat K, Hahn S (2004) Two cyclin-dependent kinases promote RNA polymerase II transcription and formation of the scaffold complex. *Mol Cell Biol* **24**: 1721-1735
- Longtine MS, McKenzie A, Demarini DJ, Shah NG, Wach A, Brachat A, Philippsen P, Pringle JR (1998) Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* **14**: 953-961
- Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**: 251-260
- Malagon F, Tong A, Shafer B, Strathern J (2004) Genetic interactions of DST1 in *Saccharomyces cerevisiae* suggest a role of TFIIS in the initiation-elongation transition. *Genetics* **166**: 1215--1227
- Malik S, Roeder RG (2010) The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nat Rev Genet* **11**: 761-772
- Mayer A, Lidschreiber M, Siebert M, Leike K, Söding J, Cramer P (2010) Uniform transitions of the general RNA polymerase II transcription complex. *Nat Struct Mol Biol* **17**: 1272-1278
- McCoy A, Grosse-Kunstleve R, Adams P, Winn M, Storoni L, Read R (2007) Phaser crystallographic software. *Journal of Applied Crystallography* **40**: 658-674
- Meinhart A, Kamenski T, Hoepfner S, Baumli S, Cramer P (2005) A structural perspective of CTD function. *Genes Dev* **19**: 1401-1415
- Mencía M, Moqtaderi Z, Geisberg JV, Kuras L, Struhl K (2002) Activator-specific recruitment of TFIID and regulation of ribosomal protein genes in yeast. *Mol Cell* **9**: 823-833
- Meyer KD, Lin S-C, Bernecky C, Gao Y, Taatjes DJ (2010) p53 activates transcription by directing structural shifts in Mediator. *Nat Struct Mol Biol* **17**: 753-760
- Miller C, Schwalb B, Maier K, Schulz D, Dümcke S, Zacher B, Mayer A, Sydow J, Marciniowski L, Dölken L, Martin DE, Tresch A, Cramer P (2011) Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast. *Mol Syst Biol* **7**: 458
- Miura F, Kawaguchi N, Sese J, Toyoda A, Hattori M, Morishita S, Ito T (2006) A large-scale full-length cDNA analysis to explore the budding yeast transcriptome. *Proc Natl Acad Sci USA* **103**: 17846-17851
- Mohibullah N, Hahn S (2008) Site-specific cross-linking of TBP in vivo and in vitro reveals a direct functional interaction with the SAGA subunit Spt3. *Genes Dev* **22**: 2994-3006
- Mooney RA, Darst SA, Landick R (2005) Sigma and RNA polymerase: an on-again, off-again relationship? *Mol Cell* **20**: 335-345

- Morris E, Ji J, Yang F, Di Stefano L, Herr A, Moon N, Kwon E, Haigis K, Naar A, Dyson N (2008) E2F1 represses beta-catenin transcription and is antagonized by both pRB and CDK8. *Nature* **455**: 552-556
- Mosley AL, Sardu ME, Pattenden SG, Workman JL, Florens L, Washburn MP (2010) Highly reproducible label free quantitative proteomic analysis of RNA polymerase complexes. *Molecular & cellular proteomics : MCP*
- Müller F, Demény MA, Tora L (2007) New problems in RNA polymerase II transcription initiation: matching the diversity of core promoters with a variety of promoter recognition factors. *J Biol Chem* **282**: 14685-14689
- Muncke N, Jung C, Rudiger H, Ulmer H, Roeth R, Hubert A, Goldmuntz E, Driscoll D, Goodship J, Schon K, Rappold G (2003) Missense mutations and gene interruption in PROSIT240, a novel TRAP240-like gene, in patients with congenital heart defect (transposition of the great arteries). *Circulation* **108**: 2843-2850
- Myers LC, Kornberg RD (2000) Mediator of transcriptional regulation. *Annu Rev Biochem* **69**: 729-749
- Naar A, Lemon B, Tjian R (2001) Transcriptional coactivator complexes. *Annu Rev Biochem* **70**: 475-501
- Näär AM, Taatjes DJ, Zhai W, Nogales E, Tjian R (2002) Human CRSP interacts with RNA polymerase II CTD and adopts a specific CTD-bound conformation. *Genes Dev* **16**: 1339-1344
- Nikolov DB, Chen H, Halay ED, Usheva AA, Hisatake K, Lee DK, Roeder RG, Burley SK (1995) Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* **377**: 119-128
- Nonet ML, Young RA (1989) Intragenic and extragenic suppressors of mutations in the heptapeptide repeat domain of *Saccharomyces cerevisiae* RNA polymerase II. *Genetics* **123**: 715-724
- Ouali M, King RD (2000) Cascaded multiple classifiers for secondary structure prediction. *Protein Sci* **9**: 1162-1176
- Pal M, Ponticelli AS, Luse DS (2005) The role of the transcription bubble and TFIIB in promoter clearance by RNA polymerase II. *Mol Cell* **19**: 101-110
- Paoletti AC, Parmely TJ, Tomomori-Sato C, Sato S, Zhu D, Conaway RC, Conaway JW, Florens L, Washburn MP (2006) Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors. *Proc Natl Acad Sci USA* **103**: 18928-18933
- Papai G, Tripathi MK, Ruhlmann C, Layer JH, Weil PA, Schultz P (2010) TFIIA and the transactivator Rap1 cooperate to commit TFIID for transcription initiation. *Nature* **465**: 956-960
- Pardee TS, Bangur CS, Ponticelli AS (1998) The N-terminal region of yeast TFIIB contains two adjacent functional domains involved in stable RNA polymerase II binding and transcription start site selection. *J Biol Chem* **273**: 17859-17864
- Pereira LA, Klejman MP, Timmers HTM (2003) Roles for BTAF1 and Mot1p in dynamics of TATA-binding protein and regulation of RNA polymerase II transcription. *Gene* **315**: 1-13
- Perrakis A, Harkiolaki M, Wilson KS, Lamzin VS (2001) ARP/wARP and molecular replacement. *Acta Crystallogr D Biol Crystallogr* **57**: 1445-1450
- Phatnani HP, Greenleaf AL (2006) Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev* **20**: 2922-2936
- Philibert R, Madan A (2007) Role of MED12 in transcription and human behavior. *Pharmacogenomics* **8**: 909-916
- Proudfoot NJ (1989) How RNA polymerase II terminates transcription in higher eukaryotes. *Trends Biochem Sci* **14**: 105-110
- Proudfoot NJ, Furger A, Dye MJ (2002) Integrating mRNA processing with transcription. *Cell* **108**: 501-512

- Puig O, Caspary F, Rigaut G, Rutz B, Bouveret E, Bragado-Nilsson E, Wilm M, Seraphin B (2001) The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods* **24**: 218-229
- Qiu H, Hu C, Zhang F, Hwang GJ, Swanson MJ, Boonchird C, Hinnebusch AG (2005) Interdependent recruitment of SAGA and Srb mediator by transcriptional activator Gcn4p. *Molecular and Cellular Biology* **25**: 3461-3474
- Raisner RM, Hartley PD, Meneghini MD, Bao MZ, Liu CL, Schreiber SL, Rando OJ, Madhani HD (2005) Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin. *Cell* **123**: 233-248
- Ranish J, Hahn S (1991) The yeast general transcription factor TFIIA is composed of two polypeptide subunits. *J Biol Chem* **266**: 19320-19327
- Ranish JA, Yudkovsky N, Hahn S (1999) Intermediates in formation and activity of the RNA polymerase II preinitiation complex: holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIIB. *Genes Dev* **13**: 49-63
- Ream TS, Haag JR, Wierzbicki AT, Nicora CD, Norbeck AD, Zhu J-K, Hagen G, Guilfoyle TJ, Pasa-Tolić L, Pikaard CS (2009) Subunit compositions of the RNA-silencing enzymes Pol IV and Pol V reveal their origins as specialized forms of RNA polymerase II. *Mol Cell* **33**: 192-203
- Reeve JN (2003) Archaeal chromatin and transcription. *Mol Microbiol* **48**: 587-598
- Reeves WM, Hahn S (2005) Targets of the Gal4 transcription activator in functional transcription complexes. *Molecular and Cellular Biology* **25**: 9092-9102
- Richard P, Manley JL (2009) Transcription termination by nuclear RNA polymerases. *Genes Dev* **23**: 1247-1269
- Rodríguez-Navarro S (2009) Insights into SAGA function during gene expression. *EMBO Rep* **10**: 843-850
- Roeder RG (1996) Nuclear RNA polymerases: role of general initiation factors and cofactors in eukaryotic transcription. *Meth Enzymol* **273**: 165-171
- Roeder RG, Rutter WJ (1969) Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature* **224**: 234-237
- Saeed A, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, Sturn A, Snuffin M, Rezantsev A, Popov D, Ryltsov A, Kostukovich E, Borisovsky I, Liu Z, Vinsavich A, Trush V, Quackenbush J (2003) TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* **34**: 374-378
- Sambrook J, Russell DW (2001) *Molecular Cloning: A Laboratory Manual*.
- Sandman K, Reeve JN (2005) Archaeal chromatin proteins: different structures but common function? *Curr Opin Microbiol* **8**: 656-661
- Segal E, Fondufe-Mittendorf Y, Chen L, Thåström A, Field Y, Moore IK, Wang J-PZ, Widom J (2006) A genomic code for nucleosome positioning. *Nature* **442**: 772-778
- Segal E, Widom J (2009) Poly(dA:dT) tracts: major determinants of nucleosome organization. *Curr Opin Struct Biol* **19**: 65-71
- Sikorski TW, Buratowski S (2009) The basal initiation machinery: beyond the general transcription factors. *Curr Opin Cell Biol* **21**: 344-351
- Singh H, Erkin A, Kremer S, Duttweiler H, Davis D, Iqbal J, Gross R, Gross D (2006) A functional module of yeast mediator that governs the dynamic range of heat-shock gene expression. *Genetics* **172**: 2169-2184

- Smale ST (1997) Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. *Biochim Biophys Acta* **1351**: 73-88
- Smyth G (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* **3**: Article3
- Söding J, Biegert A, Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic acids research* **33**: W244-248
- Soppa J (1999) Transcription initiation in Archaea: facts, factors and future aspects. *Mol Microbiol* **31**: 1295-1305
- Spahr H, Samuelsen C, Baraznenok V, Ernest I, Huylebroeck D, Remacle J, Samuelsson T, Kieselbach T, Holmberg S, Gustafsson C (2001) Analysis of Schizosaccharomyces pombe mediator reveals a set of essential subunits conserved between yeast and metazoan cells. *Proc Natl Acad Sci U S A* **98**: 11985-11990
- Stevens A (1960) Incorporation of the adenine ribonucleotide into RNA by cell fractions from E. coli B. *Biochem Biophys Res Commun* **3**: 92-96
- Struhl K (1985) Naturally occurring poly(dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. *Proc Natl Acad Sci USA* **82**: 8419-8423
- Struhl K (1989) Molecular mechanisms of transcriptional regulation in yeast. *Annu Rev Biochem* **58**: 1051-1077
- Svejstrup JQ (2004) The RNA polymerase II transcription cycle: cycling through chromatin. *Biochim Biophys Acta* **1677**: 64-73
- Taatjes DJ, Nääär AM, Andel F, Nogales E, Tjian R (2002) Structure, function, and activator-induced conformations of the CRSP coactivator. *Science* **295**: 1058-1062
- Taatjes DJ, Schneider-Poetsch T, Tjian R (2004) Distinct conformational states of nuclear receptor-bound CRSP-Med complexes. *Nat Struct Mol Biol* **11**: 664-671
- Takagi Y, Calero G, Komori H, Brown JA, Ehrensberger AH, Hudmon A, Asturias F, Kornberg RD (2006) Head module control of mediator interactions. *Mol Cell* **23**: 355-364
- Takagi Y, Kornberg R (2006) Mediator as a general transcription factor. *J Biol Chem* **281**: 80-89
- Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ (2007) How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol* **14**: 1025-1040
- Teixeira MC, Raposo LR, Mira NP, Lourenço AB, Sá-Correia I (2009) Genome-wide identification of Saccharomyces cerevisiae genes required for maximal tolerance to ethanol. *Appl Environ Microbiol* **75**: 5761-5772
- Terwilliger T, Berendzen J (1999) Automated MAD and MIR structure solution. *Acta Crystallogr D Biol Crystallogr* **55**: 849-861
- Thakur J, Arthanari H, Yang F, Pan S, Fan X, Breger J, Frueh D, Gulshan K, Li D, Mylonakis E, Struhl K, Moye-Rowley W, Cormack B, Wagner G, Naar A (2008) A nuclear receptor-like pathway regulating multidrug resistance in fungi. *Nature* **452**: 604-609
- Thomas MC, Chiang C-M (2006) The general transcription machinery and general cofactors. *Critical Reviews in Biochemistry and Molecular Biology* **41**: 105-178
- Thompson C, Koleske A, Chao D, Young R (1993) A multisubunit complex associated with the RNA polymerase II CTD and TATA-binding protein in yeast. *Cell* **73**: 1361-1375
- Thompson C, Young R (1995) General requirement for RNA polymerase II holoenzymes in vivo. *Proc Natl Acad Sci U S A* **92**: 4587-4590

- Tirosh I, Barkai N (2008) Two strategies for gene regulation by promoter nucleosomes. *Genome Res* **18**: 1084-1091
- Triezenberg SJ, Kingsbury RC, McKnight SL (1988) Functional dissection of VP16, the trans-activator of herpes simplex virus immediate early gene expression. *Genes Dev* **2**: 718-729
- Tsai FT, Sigler PB (2000) Structural basis of preinitiation complex assembly on human pol II promoters. *EMBO J* **19**: 25-36
- van de Peppel J, Kemmeren P, van Bakel H, Radonjic M, van Leenen D, Holstege FCP (2003) Monitoring global messenger RNA changes in externally controlled microarray experiments. *EMBO Rep* **4**: 387-393
- van de Peppel J, Kettelarij N, van Bakel H, Kockelkorn T, van Leenen D, Holstege F (2005) Mediator expression profiling epistasis reveals a signal transduction pathway with antagonistic submodules and highly specific downstream targets. *Mol Cell* **19**: 511-522
- van Werven FJ, van Bakel H, van Teeffelen HAAM, Altelaar AFM, Koerkamp MG, Heck AJR, Holstege FCP, Timmers HTM (2008) Cooperative action of NC2 and Mot1p to regulate TATA-binding protein function across the genome. *Genes Dev* **22**: 2359-2369
- Venters BJ, Wachi S, Mavrich TN, Andersen BE, Jena P, Sinnamon AJ, Jain P, Roller NS, Jiang C, Hemeryck-Walsh C, Pugh BF (2011) A Comprehensive Genomic Binding Map of Gene and Chromatin Regulatory Proteins in *Saccharomyces*. *Mol Cell* **41**: 480-492
- Vojnić E, Mourão A, Seizl M, Simon B, Wenzek L, Larivière L, Baumli S, Baumgart K, Meisterernst M, Sattler M, Cramer P (2011) The Mediator Med25 activator interaction domain: Structure and cooperative binding of VP16 subdomains. *Nat Struct Mol Biol*
- Wada T, Takagi T, Yamaguchi Y, Ferdous A, Imai T, Hirose S, Sugimoto S, Yano K, Hartzog GA, Winston F, Buratowski S, Handa H (1998) DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. *Genes Dev* **12**: 343-356
- Wang W, Carey M, Gralla JD (1992) Polymerase II promoter activation: closed complex formation and ATP-driven start site opening. *Science* **255**: 450-453
- Weiss S, Gladstone L (1959) A mammalian system for the incorporation of cytidine triphosphate into ribonucleic acid. *J Am Chem Soc* **81**: 4118
- Wen Y, Shatkin AJ (1999) Transcription elongation factor hSPT5 stimulates mRNA capping. *Genes Dev* **13**: 1774-1779
- West ML, Corden JL (1995) Construction and analysis of yeast RNA polymerase II CTD deletion and substitution mutations. *Genetics* **140**: 1223-1233
- Wood V, Gwilliam R, Rajandream M-A, Lyne M, Lyne R, Stewart A, Sgouros J, Peat N, Hayles J, Baker S, Basham D, Bowman S, Brooks K, Brown D, Brown S, Chillingworth T, Churcher C, Collins M, Connor R, Cronin A, Davis P, Feltwell T, Fraser A, Gentles S, Goble A, Hamlin N, Harris D, Hidalgo J, Hodgson G, Holroyd S, Hornsby T, Howarth S, Huckle EJ, Hunt S, Jagels K, James K, Jones L, Jones M, Leather S, McDonald S, McLean J, Mooney P, Moule S, Mungall K, Murphy L, Niblett D, Odell C, Oliver K, O'Neil S, Pearson D, Quail MA, Rabinowitsch E, Rutherford K, Rutter S, Saunders D, Seeger K, Sharp S, Skelton J, Simmonds M, Squares R, Squares S, Stevens K, Taylor K, Taylor RG, Tivey A, Walsh S, Warren T, Whitehead S, Woodward J, Volckaert G, Aert R, Robben J, Grymonprez B, Weltjens I, Vanstreels E, Rieger M, Schäfer M, Müller-Auer S, Gabel C, Fuchs M, Düsterhöft A, Fritzc C, Holzer E, Moestl D, Hilbert H, Borzym K, Langer I, Beck A, Lehrach H, Reinhardt R, Pohl TM, Eger P, Zimmermann W, Wedler H, Wambutt R, Purnelle B, Goffeau A, Cadieu E, Dréano S, Gloux S, Lelaure V, Mottier S, Galibert F, Aves SJ, Xiang Z, Hunt C, Moore K, Hurst SM, Lucas M, Rochet M, Gaillardin C, Tallada VA, Garzon A, Thode G, Daga RR, Cruzado L, Jimenez J, Sánchez M, del Rey F, Benito J, Domínguez A, Revuelta JL, Moreno S, Armstrong J, Forsburg SL, Cerutti L, Lowe T, McCombie WR, Paulsen I, Potashkin J, Shpakovski GV, Ussery D, Barrell BG, Nurse P, Cerrutti L (2002) The genome sequence of *Schizosaccharomyces pombe*. *Nature* **415**: 871-880
- Wu Z, Irizarry R, Gentleman R, Martinez-Murillo F, Spencer F (2004) A Model-Based Background Adjustment for Oligonucleotide Expression Arrays. *J Am Stat Assoc* **99**: 909-917

- Xu X, Zhou H, Boyer TG (2011) Mediator is a transducer of amyloid-precursor-protein-dependent nuclear signalling. *EMBO Rep*
- Yamaguchi Y, Takagi T, Wada T, Yano K, Furuya A, Sugimoto S, Hasegawa J, Handa H (1999) NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell* **97**: 41-51
- Yang F, Vought B, Satterlee J, Walker A, Jim Sun Z, Watts J, DeBeaumont R, Saito R, Hyberts S, Yang S, Macol C, Iyer L, Tjian R, van den Heuvel S, Hart A, Wagner G, Naar A (2006) An ARC/Mediator subunit required for SREBP control of cholesterol and lipid homeostasis. *Nature* **442**: 700-704
- Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J, Speed TP (2002) Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic acids research* **30**: e15
- Yuan G-C, Liu Y-J, Dion MF, Slack MD, Wu LF, Altschuler SJ, Rando OJ (2005) Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science* **309**: 626-630
- Yudkovsky N, Ranish JA, Hahn S (2000) A transcription reinitiation intermediate that is stabilized by activator. *Nature* **408**: 225-229
- Yuryev A, Corden JL (1996) Suppression analysis reveals a functional difference between the serines in positions two and five in the consensus sequence of the C-terminal domain of yeast RNA polymerase II. *Genetics* **143**: 661-671
- Zhang H, Roberts DN, Cairns BR (2005) Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss. *Cell* **123**: 219-231
- Zhang Z, Reese JC (2004) Redundant mechanisms are used by Ssn6-Tup1 in repressing chromosomal gene transcription in *Saccharomyces cerevisiae*. *J Biol Chem* **279**: 39240-39250
- Zhu W, Smith JW, Huang C-M (2010) Mass spectrometry-based label-free quantitative proteomics. *J Biomed Biotechnol* **2010**: 840518
- Zhu Y, Qi C, Jain S, Le Beau M, Espinosa R, Atkins G, Lazar M, Yeldandi A, Rao M, Reddy J (1999) Amplification and overexpression of peroxisome proliferator-activated receptor binding protein (PBP/PPARBP) gene in breast cancer. *Proc Natl Acad Sci U S A* **96**: 10848-10853

## Abbreviations

activator	gene-specific activator protein
AH	activation helix
bp	base pairs
BRE	TFIIB recognition element
BSA	bovine serum albumine
CTD	carboxy-terminal domain of Rpb1
CPE	core promoter element
CV	column volumes
Da	dalton DMSO dimethyl sulfoxide
DTT	1,4-dithio-D,L-threitol
<i>E. coli</i>	Escherichia coli
EM	electron microscopy
GTF	general transcription factor
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethane sulfonic acid
<i>Hs</i>	<i>Homo sapiens</i>
INR	Initiator (core promoter element)
IPTG	Isopropyl- $\beta$ -D-thiogalactopyranoside
kDa	kilo Dalton
MDa	mega Dalton
LB	Luria-Bertani (media)
MCS	multiple cloning site
MES	2-N-morpholino-ethanesulfonic acid
MOPS	4-morpholine-propanesulfonic acid
mRNA	messenger RNA
MW	molecular weight
NE	nuclear extract
OD <sub>600</sub>	optical density at a wavelength of 600 nm
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PDB	Protein Data Bank
PIC	pre-initiation complex

RNAP	DNA-dependent RNA polymerase
rpm	rounds per minute
SAGA	Spt-Ada-Gcn5 acetylase complex
Sc	<i>Saccharomyces cerevisiae</i>
Sp	<i>Schizosaccharomyces pombe</i>
SDS	sodium dodecylsulfate
TAD	transcription activation domain
TAF	TBP-associated factor
TAP	Tandem-affinity purification
TBP	TATA-binding protein
Tris	Tris-(hydroxymethyl)-aminomethane
TSS	transcription start site
u	unit
UAS	upstream activating sequence
v/v	volume per volume
w/v	weight per volume
wt	wild-type

## List of Figures

Figure 1: The central dogma of of molecular biology	1
Figure 2: RNA polymerase II transcription cycle	3
Figure 3: Regulatory factors and the basal transcription machinery at the core promoter	4
Figure 4: Med8C/18/20 is a subcomplex of Mediator head.	49
Figure 5: Structural conservation of the Med8C/18 interaction	51
Figure 6: Med8C/18/20 is required for activated transcription <i>in vitro</i> .	52
Figure 7: Comparative gene expression profiling	54
Figure 8: Structure of Med11/22 Mediator subcomplex	59
Figure 9: <i>In vivo</i> phenotyping of Med11/22 truncations	61
Figure 10: A heterodimeric four-helix bundle building block in Mediator	62
Figure 11: C-terminal helices in Med11/22 are conserved across species	63
Figure 12: C-terminal extensions of Med11/22 bind a Med17 C-terminal domain	64
Figure 13: Surface of Med11/22 helix bundle.	66
Figure 14: A conserved interaction patch on Med11/22	67
Figure 15: Med11/22 is a functional submodule regulating a specific subset of genes	68
Figure 16: Med11/22 is required for stable PIC formation <i>in vitro</i> and <i>in vivo</i>	70
Figure 17: Submodular architecture of Mediator head contacting PIC components	72
Figure 18: Recombinant <i>Sc</i> Mediator head module is functionally active	73
Figure 19: PIC formation requires the Mediator Head module	75
Figure 20: Localization of proteins identified by tandem mass spectrometry	76
Figure 21: Analysis of immobilized template assays by tandem mass spectrometry and spectral counting	79
Figure 22: Med7N/31 is a functional submodule <i>in vitro</i>	83
Figure 23: Med7N/31 cooperates with TFIIS during activated transcription <i>in vitro</i>	84
Figure 24: NMR structure of Med25 and interaction with VP16 TAD	86
Figure 25: ACID-VP16 interaction assay	87
Figure 26: VP16 activation is quenched by ACID in yeast and mammalian system	88
Figure 27: VP16 subdomains H1 and H2 activate yeast transcription synergistically	89
Figure 28: Models of the Pol II closed and open complex	90
Figure 29: B-reader and transcription start site selection	92

## List of tables

Table 1: Subunit composition of multisubunit RNAP from all three kingdoms of life	2
Table 2: Factors involved in Pol II transcription initiation	8
Table 3: Mediator subunit composition and modular architecture	12
Table 4: Yeast strains ( <i>Sc</i> )	18
Table 5: Yeast strains ( <i>Sp</i> )	19
Table 6: <i>E. coli</i> strains	19
Table 7: Plasmids used for recombinant expression of <i>Med11/22</i> in <i>E. coli</i>	19
Table 8: Plasmids used for recombinant expression of other proteins in <i>E. coli</i>	21
Table 9: Plasmids used as templates for yeast complementation experiments	22
Table 10: Plasmids used for <i>in vitro</i> assays	24
Table 11: Oligonucleotides used for <i>in vitro</i> experiments	25
Table 12: Oligonucleotides used for quantitative real-time PCR	25
Table 13: Antibodies used in this study	26
Table 14: Growth media	26
Table 15: Growth media additives	27
Table 16: General buffers and solutions	27
Table 17: Recombinant protein purification buffers	28
Table 18: Buffers for <i>in vitro</i> biochemical assays	29
Table 19: Data collection and refinement statistics for the <i>Sp Med8C/18</i> structure	50
Table 20: Data collection and refinement statistics for <i>Med11/22</i> structure	60
Table 21: Mediator and Pol II subunits detected by tandem mass spectrometry analysis of wild-type immobilized template eluates	77

## Curriculum vitae

---

### Martin Josef Seizl

Date of birth: July 8, 1981  
Place of birth: Munich, Germany  
Nationality: German

### Education

---

- Since 05/11      **PostDoc**  
University of Munich (LMU) – Gene Center, Munich, Germany  
Prof. Patrick Cramer
- 05/07 – 04/11      **PhD thesis**  
University of Munich (LMU) – Gene Center, Munich, Germany  
Prof. Patrick Cramer  
*“The Mediator head module and regulation of RNA polymerase II transcription initiation”*
- 11/06 – 04/07      **Master’s thesis**  
Fred Hutchinson Cancer Research Center, Seattle, USA  
University of Munich (LMU) – Gene Center, Munich, Germany  
Prof. Steve Hahn & Prof. Patrick Cramer  
*“A structure-function analysis of the transcription coactivator complex Mediator”*
- 10/05 – 04/07      **Biochemistry (M.Sc.)**  
University of Munich (LMU), Munich, Germany
- 04/05 – 09/05      **Industry research position**  
Crelux GmbH, Martinsried, Germany
- 04/03 – 07/03      **Bachelor’s thesis**  
Technical University of Munich (TUM) – Institute of Microbiology,  
Weihenstephan, Germany - Prof. Siegfried Scherer  
*“Degradation of PrP-res by proteases of coryneform bacteria”*
- 10/01 – 03/05      **Human nutrition (B.Sc.)**  
Technical University of Munich (TUM), Weihenstephan, Germany  
Oklahoma State University, Stillwater, USA
- 07/00 – 04/01      **German military service**
- 09/91 – 06/00      **Abitur**  
Gymnasium (Secondary School), Kirchheim, Germany

## **Publications**

---

Seizl M<sup>1</sup>, Lariviere L<sup>1</sup>, Pfaffeneder T, Wenzek L, Cramer P. A conserved four-helix bundle in Mediator subunits Med11/Med22 is required for stable preinitiation complex formation. *Nucleic Acids Res.* 2011 in press

<sup>1</sup> these authors contributed equally

Chanarat S, Seizl M, Sträßler K. The Prp19 complex is a novel transcription elongation factor. *Genes Dev.* 2011 in press

Czeko E, Seizl M, Mielke T, Cramer P. Iwr1 directs RNA polymerase II nuclear import. *Mol Cell.* 2011 in press

Vojnić E, Mourão A, Seizl M, Simon B, Wenzek L, Larivière L, Baumli S, Baumgart K, Meisterernst M, Sattler M, Cramer P. The Mediator Med25 activator interaction domain: Structure and cooperative binding of VP16 subdomains. *Nat Struct Mol Biol.* 2011 Apr;18(4):404-9. Epub 2011 Mar 6

Kostrewa D<sup>1</sup>, Zeller ME<sup>1</sup>, Armache KJ<sup>1</sup>, Seizl M, Leike K, Thomm M, Cramer P. RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature.* 2009 Nov 19;462(7271):323-30

<sup>1</sup> these authors contributed equally

Koschubs T, Seizl M, Larivière L, Kurth F, Baumli S, Martin DE, Cramer P. Identification, structure, and functional requirement of the Mediator submodule Med7N/31. *EMBO J.* 2009 Jan 7;28(1):69-80. Epub 2008 Dec 4

Lariviere L<sup>1</sup>, Seizl M<sup>1</sup>, van Wageningen S, Roether S, Feldmann H, Straesser K, Hahn S, Holstege F, Cramer P. Structure-system correlation identifies a gene regulatory Mediator submodule. *Genes Dev.* 2008 Apr 1;22(7):872-877

<sup>1</sup> these authors contributed equally

## **Fellowships & Awards**

---

since 02/08      **Elite Network of Bavaria – Graduate school**  
Program: “Protein Dynamics in Health and Disease”

08/07 – 04/10      **Boehringer Ingelheim Fonds – PhD fellowship**

04/07      **Roemer price for young scientists**  
Award for Master’s Thesis

11/06 – 12/06      **EMBO short-term fellowship**  
Support for Master’s Thesis in Seattle, USA