

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

Structure of the MED7/MED21 heterodimer and reconstitution of a recombinant Mediator middle module complex



Sonja Baumli
aus Rain, Schweiz
2005

Erklärung

Diese Dissertation wurde im Sinne von §13 Abs. 3 bzw. 4 der Promotionsordnung vom 29. Januar 1998 von Herrn Prof. Dr. Patrick Cramer betreut.

Ehrenwörtliche Versicherung

Diese Dissertation wurde selbständig und ohne unerlaubte Hilfe erarbeitet.

München, den 20.09.2005

Sonja Baumli

Dissertation eingereicht am: 21.09.05

1. Gutachter: Prof. Dr. Patrick Cramer

2. Gutachter: Dr. Michael Meisterernst

Mündliche Prüfung am: 22.11.05

Acknowledgements

I would like to thank Patrick Cramer for giving me the opportunity to do this work in his lab and for giving me the possibility to work as independently as possible. I am very grateful for the excellent working conditions and the very good atmosphere in the lab.

I am especially grateful to Sabine for all the contributions and the collaboration over the last 3 years.

I would like to thank Laurent and Michela for discussions and critical comments on this thesis. Erika, Karim, Hubert and all the other members of the Cramer Lab, thank you for the grate atmosphere, help and advice during various stages of the work.

I would like to thank Claudia for everyday help and advice in the lab and Toni for teaching me about crystallography.

I am also grateful for the support I got from Conny, Markus, Arne, Christina and Martin who did their practica and bachelor work in the lab.

I have to thank Michael Meisterernst and Uwe Jacob for very helpful advice as my PhD committee.

Thank you Tom for the constant support, it made everyday life much easier and nicer.

Summary

The Mediator of transcriptional regulation is the central coactivator that enables a response of RNA polymerase II (Pol II) to activators and repressors. Yeast Mediator has a size of more than one MDa and consists of 25 different polypeptides. Biochemical studies defined three Mediator modules in yeast, the head (MED17) the middle (MED9/MED10) and the tail (MED15) modules. During this work, an *E.coli* coexpression-copurification system was developed, which allowed to study pairwise interactions of Mediator middle module subunits. With the help of this system I reconstituted a complex of two essential and conserved yeast Mediator middle module proteins, the MED7/MED21 heterodimer, and solved its crystal structure. The heterodimer forms an extended structure, which spans one third of the Mediator length, and almost the diameter of Pol II. It shows a four helix bundle and a coiled-coil protrusion connected by a flexible hinge. Multiple conserved patches can be identified on the surface, which allow for assembly of the middle module. A combination of the coexpression-copurification system and assembly of subcomplexes allowed the reconstitution of a five-subunit Mediator middle module subcomplex. The reconstituted subcomplex is able to bind Pol II *in vitro*. MED6 associates with the middle module and forms a bridge to the head module. The potential flexibility of this bridge and the MED7/MED21 hinge can account for changes in Mediator structure upon its binding to Pol II or to activators.

Publications

Current list of publications to which this work contributed:

Baumli S., Hoepfner S., Cramer P. (2005)

A Conserved Mediator Hinge Revealed in the Structure of the MED7•MED21 (Med7•Srb7) Heterodimer.

The Journal of Biological Chemistry 280 (18), 18171-18178

Meinhart, A., Kamenski, T., Hoepfner, S., Baumli, S., and Cramer, P. (2005).

A Structural Perspective of CTD Function.

Genes and Development 19, 1401-1415.

Hoepfner S., Baumli S., Cramer P. (2005)

Structure of the Mediator Subunit Cyclin C and its Implications for CDK8 Function.

The Journal of Molecular Biology 350, 833-842.

Contents

1	Introduction	5
1.1	RNA polymerases	5
1.2	The RNA polymerase II transcription cycle	6
1.2.1	The general transcription factors and transcription initiation	7
1.2.2	The C-terminal domain of RNA polymerase II is modified during the transcription cycle	8
1.3	Regulated transcription	9
1.4	Mediator and transcription regulation	10
1.4.1	Mediator complexes have a modular architecture	11
1.4.2	The role of Mediator during transcription initiation and reinitiation	13
1.4.3	Structural information on Mediator complexes	13
1.5	The yeast Mediator proteins	15
1.5.1	The Mediator head module	15
1.5.2	The Mediator middle module	16
1.5.3	The Mediator tail module	17
1.5.4	The CDK8/CyclinC module of Mediator	17
1.5.5	MED25, an additional human Mediator subunit	18
1.6	VP16, a viral activator of transcription	19
1.7	Aims and scope of this work	20
2	Results	21
2.1	The MED7/MED21 subcomplex	21
2.1.1	Subcomplex domain mapping	21
2.1.2	Purification of crystallizable MED7/MED21 variants	23
2.1.3	Crystallization of the MED7/MED21 heterodimer	25

2.1.4	Methionine mutations enhance the anomalous signal	26
2.1.5	Structure determination of the MED7/MED21 heterodimer	27
2.1.6	The MED7/MED21 structure	28
2.1.7	Conserved interactions of the MED7/MED21 heterodimer within the Mediator middle module	34
2.1.8	MED4 and MED9	35
2.1.9	MED6 bridges two Mediator modules	36
2.1.10	MED31: an integral part of the middle module	37
2.1.11	Reconstitution of larger subcomplexes of the Mediator middle module	38
2.2	Reconstitution of a hexameric Mediator middle module	44
2.2.1	Purification of a tetrameric MED7/MED10/MED21/MED31 Mediator middle module complex:	44
2.2.2	Assembly of the hexameric Mediator middle module consist- ing of MED4/MED7/MED9/MED10/MED21/MED31	46
2.2.3	The Mediator middle module binds Pol II	52
2.3	1.3 Structural studies of MED25	55
2.3.1	Purification of ACID	55
2.3.2	ACID interacts with DNA	57
2.3.3	Mapping of a minimal VP16 activation domain	58
2.3.4	C-terminal truncation of <i>D. melanogaster</i> ACID prevents in- teraction with the VP16 AD	60
2.3.5	Crystallization trials	61
3	Discussion	64
3.1	Pairwise interactions within the Mediator middle module	64
3.2	The MED7/MED21 structure	65

3.2.1	Conserved Mediator core architecture	67
3.2.2	Conserved hinges and Mediator function	68
3.3	Larger complexes	69
3.3.1	A reconstituted Mediator middle module	70
3.4	Conservation within Mediator subunits	71
3.5	Towards the structure of an activator interaction domain	73
3.6	Conclusion and Outlook	75
4	Experimental procedures	76
4.1	General Methods	76
4.1.1	Bacterial strains	76
4.1.2	Plasmids	76
4.1.3	Growth media	83
4.1.4	Bioinformatic tools	84
4.1.5	Protein expression and selenomethionine labeling	84
4.1.6	Limited proteolysis experiments	85
4.1.7	Protein Analysis	85
4.1.8	Gelfiltration experiments and calculation of the hydrodynamic radius from the structure	86
4.1.9	Static light scattering	86
4.2	The middle module	87
4.2.1	Strategy for the expression of the middle module proteins	87
4.2.2	Cloning	87
4.2.3	Expression and purification of middle module proteins	88
4.2.4	Interaction assay	90
4.2.5	Crystallization and crystal treatment	91
4.2.6	X-ray structure determination	92

4.3	Reconstitution of the middle module	94
4.3.1	Expression and purification of the MED7/MED10/MED21/MED30 middle module complex	94
4.3.2	Reconstitution of a MED4/MED7/MED9/MED10/MED21/MED30 middle module complex	95
4.3.3	Assembly of a Pol II/middle module complex	95
4.3.4	Gst-CTD pull-down	96
4.4	ACID	97
4.4.1	Design of ACID variants	97
4.4.2	Purification of ACID constructs	97
4.4.3	Purification of the ACID/VP16 complex	100
4.4.4	Binding of ACID to Gst-VP16	100
4.4.5	Electrophoretic mobility shift assay	101
4.4.6	Crystallization of ACID variants	102

1 Introduction

1.1 RNA polymerases

Transcription is the fundamental process of RNA synthesis from a DNA template. The basic transcription reaction is very similar in all organisms. RNA polymerases are the enzymes which transcribe DNA into RNA. Sequence and structural comparisons reveal the evolutionary conservation of the overall architecture of RNA polymerases in the three kingdoms of life. The five subunits of the bacterial polymerase form a core homologous to the yeast RNA polymerase core (reviewed in Cramer, 2002). Archaeal and eukaryotic RNA polymerases are multi-subunit enzymes, contain at least ten subunits and are highly homologous. While bacteria and archaea have only one RNA polymerase, eukaryotes have three enzymes to synthesize different classes of RNA. RNA polymerase I (Pol I) is mainly responsible for the synthesis of ribosomal RNA (rRNA), RNA polymerase II (Pol II) transcribes protein coding genes into precursor mRNA, and RNA polymerase III (Pol III) transcribes one rRNA and all transfer RNAs. Bacterial, archaeal and eukaryotic polymerases have their own set of transcription factors. In bacteria the most prominent accessory factor is the $\sigma 70$ factor. It recognizes the promoter sequences of house keeping genes, promotes conformational changes upon initiation and interacts directly with regulatory proteins (reviewed in Murakami and Darst, 2003). In archaea and eukaryotes the function of the σ subunit has been replaced by an abundance of transcription factors. The archaeal polymerase relies on two essential general transcription factors, the TATA binding protein (TBP) and TFB. Transcriptional regulators seem to be adopted largely from bacterial systems and large coactivator complexes are absent (reviewed in Geiduschek and Ouhammouch, 2005; Ouhammouch, 2004). Eukaryotic Pol II has been extensively characterized, both biochemically and structurally. Transcription of protein-coding genes by Pol II is dependent on additional factors. Pol II associates with the five general transcription factors TFIIA, -B, -D, -E, -H, coactivators, mRNA processing factors and RNA export factors (Woychik and Hampsey, 2002; Hahn, 2004). The crystal structures of Pol II, Pol II with nucleic acids and Pol II with the general transcription factors TFIIS and TFIIB have been determined (Cramer et al., 2001; Armache et al., 2005; Bushnell et al., 2004; Kettenberger et al., 2004). EM reconstructions of a Pol II/TFIIF complex have also been achieved (Chung et al., 2003; Asturias, 2004).

1.2 The RNA polymerase II transcription cycle

Transcription catalyzed by Pol II involves multiple steps. The transcription cycle can be divided in initiation, elongation, termination and reinitiation. During initiation, gene specific regulatory factors recruit chromatin remodeling factors or coactivators that interact directly with the transcriptional machinery to promote preinitiation complex (PIC) assembly. Coactivators, such as Mediator, transmit regulatory signals to the polymerase. General transcription factors TFIIB, -D, -E, -F, and H mediate promoter recognition and opening. Initially, in the PIC all factors and the polymerase are bound to the promoter in a so-called “closed” complex. Transition to the “open” complex involves DNA melting, requires ATP hydrolysis and in the presence of NTPs the polymerase initiates transcription. After the synthesis of about 30 nucleotides, the polymerase leaves the initiation complex and enters the elongation phase (promoter clearance). Elongation factors ensure efficient mRNA production. Chromatin remodeling factors, RNA processing and polyadenylation factors are recruited to the polymerase. Most of the factors which are involved in initiation remain bound to the promoter in a scaffold complex (Yudkovsky et al., 2000). The scaffold complex is thought to promote fast transcription reinitiation at a previously transcribed gene (Figure 1; Hahn, 2004).

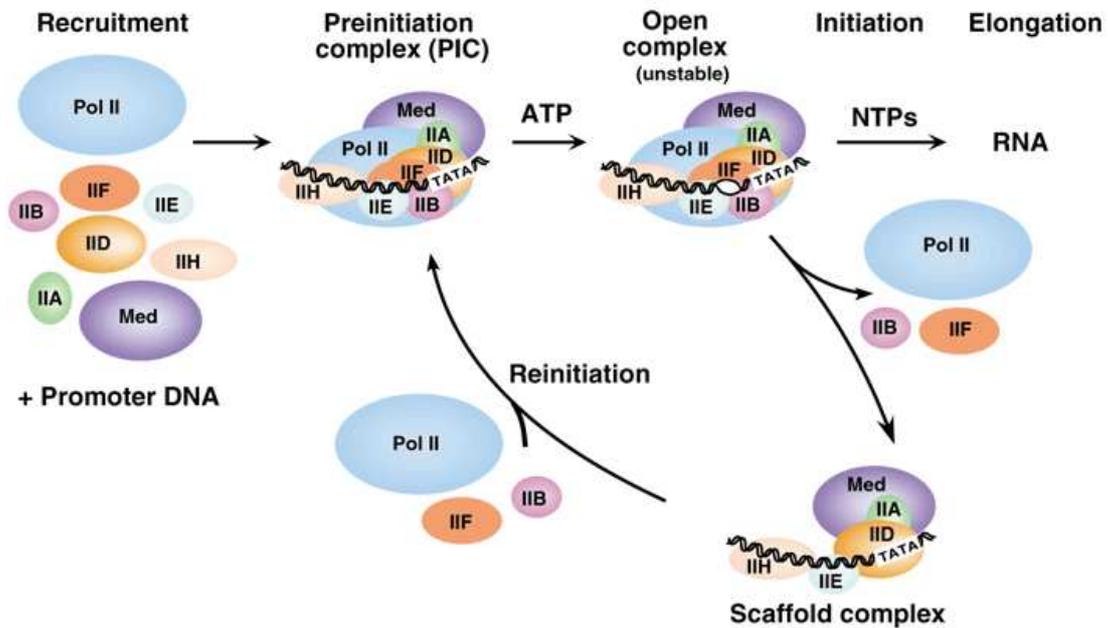


Figure 1: Transcription initiation and reinitiation pathways for Pol II. (adopted from Hahn, 2004)

1.2.1 The general transcription factors and transcription initiation

During transcription initiation the basal transcription machinery is assembled at the core promoter, the minimal sequence which is required for basal transcription. The core promoter includes defined sequence elements within 35 bp upstream the transcription start site in human and drosophila, 40 to 120 bp in yeast. These elements include the TATA element (TBP binding), TFIIB recognition element (BRE), Initiator element (INR) and Downstream promoter element (DPE) (reviewed in Smale and Kadonaga, 2003). The most prominent core promoter sequence is the TATA element which is bound by TBP. TBP binding promotes DNA bending and seems to be the first step in transcription machinery assembly (Kim et al., 1993). TBP belongs to the 15 subunit TFIID complex. TFIIA binding stabilizes the TBP/DNA complex (Geiger et al., 1996; Tan et al., 1996). TFIIB directly contacts TBP as well as the DNA in the complex (Nikolov et al., 1995). Polymerase and TFIIF can then assemble with the preformed complex. After the recruitment of TFIIE and TFIIH, ATP dependent promoter opening can occur. The helicase of TFIIH is responsible for open complex formation. During this process the DNA template strand is positioned within the polymerase active site. TFIIH also phosphorylates the C-terminal domain (CTD) of the largest subunit

Factor	Function
TFIIA	Stabilizes TBP and TFIID-DNA binding. Blocks transcription inhibitors. Positive and negative gene regulation.
TFIIB	Binds TBP, Pol II and promoter DNA. Helps fix the transcription start site.
TFIID	TBP Binds TATA element and deforms promoter DNA. Platform for assembly of TFIIB, TFIIA and TAFs.
	TAFs Binds INR and DPE elements. Target of regulatory factors.
Mediator	Binds cooperatively with Pol II. Kinase and acetyl transferase activity. Stimulates basal and activated transcription. Target of regulatory factors.
TFIIF	Binds Pol II and is involved in Pol II recruitment to PIC and in open complex formation.
TFIIE	Binds promoter near transcription start.
TFIIH	Functions in transcription and DNA repair. Kinase and two helicase activities. Essential for open complex formation.
SAGA	Contains Histone acetyl-transferase and Ubiquitin protease activity. Interacts with TBP and TFIIA. Target of activators.

Table 1: General transcription factors and coactivators in the yeast *S. cerevisiae*. (adapted from Hahn, 2004)

of Pol II thereby enabling polymerase dissociation from the PIC (Table 1). Basal transcription has been reconstituted *in vitro* by adding general transcription factors and polymerase to the DNA template (Sayre et al., 1992).

1.2.2 The C-terminal domain of RNA polymerase II is modified during the transcription cycle

Pol II Rbp1 contains a CTD which consists of 25 to 52 heptad repeats, according to the organism, with the consensus sequence YSPTSPS (Corden, 1990). This consensus sequence acts as a platform for assembly of factors that regulate transcription initiation, elongation, termination and mRNA processing. CTD phosphorylation or dephosphorylation coordinates the transcription cycle (Dahmus, 1996). During initiation, Mediator binding to Pol II requires a dephosphorylated CTD. CTD phosphorylation breaks Pol II-Mediator interactions, resulting in an elongating Pol II and a scaffold complex bound to the promoter. The elongating Pol II has a phosphorylated CTD. The CDK/Cyclin pairs of TFIIH (CDK7/CyclinH) and Mediator (CDK8/CyclinC) phosphorylate the CTD during initiation. The CDK/Cyclin pair of pTEFb (CDK9/CyclinT)

phosphorylates the CTD during elongation (Pinhero et al., 2004). Proteins required for mRNA processing bind to the phosphorylated CTD. The phosphatases Fcp1, Scp1 (Lin et al., 2002; Kamenski et al., 2004) and Ssu72 (Meinhart et al., 2003b) dephosphorylate the CTD and thereby enable the interaction with the scaffold complex and transcription reinitiation.

The CTD, together with a linker region, is disordered in Pol II crystal structures (Cramer et al., 2001). NMR and CD studies also show a largely unstructured nature of the free CTD (reviewed in Meinhart et al., 2005). The CTD could provide a platform for many different interactions with structurally different target proteins via an *induced fit* mechanism (Lima, 2005). Only peptide structures of the phosphorylated CTD in complex with Pcf11 (Meinhart and Cramer, 2004), Ctg1 (Fabrega et al., 2003) and Pin1 (Verdecia et al., 2000) have been determined. The CTD in the three structures is differentially modified and shows different conformations (reviewed in Meinhart et al., 2005; Fabrega et al., 2003).

The non-phosphorylated CTD is required for Mediator binding. In yeast the CTD is necessary for the formation of a stable Pol II/Mediator complex (Myers et al., 1998; Asturias et al., 1999), and CTD-Mediator interaction is needed for Mediator function, since yeast Mediator cannot stimulate transcription by a CTD-less Pol II. The human Mediator Complex CRSP (cofactor required for Sp1) also interacts with the CTD and adopts a specific CTD bound conformation (Naar et al., 2002). However, there is no indication on how the non-phosphorylated CTD might interact with Mediator. Larger isoforms of Mediator such as the Mediator-like complexes NAT (negative regulator of activated transcription) and SMCC (Srb/Med-containing cofactor complex) can act independently of the CTD (Sun et al., 1998; Gu et al., 1999). Both Complexes include the CTD kinase of the CDK8/CyclinC Module.

1.3 Regulated transcription

Regulation of expression of protein-coding genes is very important during cell growth and differentiation. Gene expression regulation occurs at many different levels, including transcription, RNA processing, mRNA export, translation as well as mRNA and protein stability. A very strongly regulated process however, is transcription initiation.

Regulation of bacterial transcription requires transcriptional activators or repressors which bind to DNA elements and directly exert their function on the bacterial RNA

polymerase. Transcription regulation in eukaryotes is more elaborate, since the organisms are more complex (Levine and Tjian, 2003). Multiple regulatory signals which come from DNA bound, gene specific activators or repressors have to be processed and transmitted to the Pol II transcription machinery. These gene specific transcription regulators recruit multi-protein coactivators, *via* direct protein-protein interactions (reviewed in Kadonaga, 2004; Naar et al., 2001). Coactivators can act indirectly or directly to regulate the activity of the Pol II transcription machinery at the core promoter. Chromatin related coactivators regulate transcription indirectly by nucleosome remodeling or by covalent Histone modifications. Within the Pol II transcriptional machinery, Mediator and TFIID are coactivators which are targets of transcriptional regulators. These two coactivators regulate transcription directly, by forming an interface between regulators and Pol II in complex with the general transcription factors. Over the last decade evidence from many laboratories has converged on Mediator as the central Pol II coactivator (Myers and Kornberg, 2000; Malik and Roeder, 2000; Bjorklund and Gustafsson, 2004; Kornberg, 2005; Malik and Roeder, 2005).

1.4 Mediator and transcription regulation

Mediator was discovered in the yeast *S. cerevisiae* by its ability to enable activated transcription in an *in vitro* system containing Pol II and the general transcription factors (Flanagan et al., 1991; Kelleher et al., 1990; Kim et al., 1994). Human Mediator was described in a screen for transcription cofactors (Meisterernst et al., 1991; Kretschmar et al., 1994). Genome-wide studies showed that Mediator is required for regulated transcription of the majority of yeast genes (Holstege et al., 1998). Mediator promotes initiation complex assembly through activator-Mediator, Mediator-Pol II and Mediator-general transcription factor contacts (Cantin et al., 2003). Initial attempts to purify yeast Mediator yielded a “holoenzyme”, consisting of Mediator in complex with Pol II and loosely associated factors of the basal transcription machinery (Hengartner et al., 1995; Wilson et al., 1996). These findings lead to the proposition that it is a “holoenzyme” which is recruited to the promoter upon gene induction (Kim et al., 1994). Later studies on the *in vivo* distribution of Pol II and Mediator on *Drosophila* polytene chromosomes (Park et al., 2001b) showed that the two complexes do not colocalize. Recruitment of Mediator precedes the recruitment of Pol II (Bhoite et al., 2001; Cosma et al., 2001). Accordingly, about 70 % of Mediator in yeast occurs as

free Mediator and only 30 to 40 % as Mediator/Pol II complex (Liu et al., 2001; Takagi et al., 2005). Mediator also stimulates basal transcription (Baek et al., 2002; Mittler et al., 2001). The yeast Mediator comprises 25 polypeptide subunits: 11 subunits are essential and 22 are at least partially conserved in sequence throughout eukaryotes (Bourbon et al., 2004; Boube et al., 2002). Mediator-like complexes were isolated from human cells (Fondell et al., 1996; Sun et al., 1998; Naar et al., 1999; Rachez et al., 1999; Ryu and Tjian, 1999), mouse (Jiang et al., 1998), fission yeast (Spahr et al., 2000) and *Drosophila* (Park et al., 2001a). Comparison of the different Mediators has led to the identification of homologues of most of the yeast Mediator subunits (Boube et al., 2002; Bourbon et al., 2004; Table 2).

1.4.1 Mediator complexes have a modular architecture

The Mediator subunit architecture was inferred from biochemical, genetic, and electron microscopic studies. Biochemical studies defined three Mediator submodules, which were stable even in the presence of elevated urea concentrations: the head module (MED17/Srb4), the middle module (MED9/MED10) and the tail module (Rgr1/MED14), which were tentatively correlated with three density lobes in electron microscopic images. The head, middle and tail modules are found in most Mediator preparations. An additional module, the CDK8/CyclinC module is only partially associated to the Mediator (Samuelsen et al., 2003). The yeast Mediator has been purified as a single multisubunit complex. On the contrary, different purifications of mammalian Mediator complexes often have different compositions. Purification schemes based on interactions with transcriptional activators (TRAP (Fondell et al., 1996), ARC (Naar et al., 1999), DRIP (Rachez et al., 1999), E1A (Boyer et al., 1999)), homologous subunits to yeast Mediator (SMCC (Gu et al., 1999), NAT (Sun et al., 1998)) and coactivator activity (CRSP (Ryu and Tjian), PC2 (Malik et al., 2000)) purified complexes of different subunit compositions. Even though there are differences in the preparations, all complexes contain a core of subunits which have orthologous in *S. cerevisiae* Mediator (Taatjes et al., 2004a). In addition, proteomic analysis of affinity-purified Mediator, with the affinity tag on different subunits, resulted in one complex with all identified subunits present (Sato et al., 2004). These observations indicate that the purified complexes probably represent versions of the same cellular entity (Conaway et al., 2005).

new Name	<i>S. cerevisiae</i> (previous name)	<i>H. sapiens</i> (previous name)	location
MED1	Med1	TRAP220	middle
MED2	Med2	-	tail
MED3	Pgd1/Hrs1/Med3	-	tail
MED4	Med4	TRAP36/DRIP36	middle
MED5	Nut1	-	middle/tail
MED6	Med6	hMed6/Drip33	head
MED7	Med7	hMed7/Drip34	middle
MED8	Med8	Arc32	head
MED9	Cse2/Med9	Med25	middle
MED10	Nut2/Med10	hNut2/hMed10	middle
MED11	Med11	HSPC296	head
MED12	Srb8	TRAP230/DRIP240	kinase
MED13	Srb9	TRAP240/DRIP250	kinase
MED14	Rgr1	TRAP170/DRIP150/CRSP150	tail
MED15	Gal11	ARC105	tail
MED16	Sin4	TRAP95/DRIP92	tail
MED17	Srb4	TRAP80/DRIP77/CRSP77	head
MED18	Srb5	p28b	head
MED19	Rox3	LCMR1	head
MED20	Srb2	hTRFP/p28a	head
MED21	Srb7	hSrb7/p21	middle
MED22	Srb6	Med24/Surf5	head
MED23	-	TRAP150 β /DRIP130/CRSP130/hSur2	?
MED24	-	TRAP100/DRIP100/CRSP100	?
MED25	-	ARC92/ACID1	?
MED26	-	ARC70/CRSP70	?
MED27	-	TRAP37/CRSP347	?
MED28	-	Fksg20	?
MED29	-	Hintersex	?
MED30	-	TRAP25	?
MED31	Soh1	hSoh1	middle
CDK8	Srb10/Ssn3/Ume5	hSrb10/CDK8	kinase
Cyclin C	Srb11/Ssn8/Ume3	hSrb11/CycC	kinase

Table 2: Mediator subunit nomenclature according to Bourbon et al., 2004.

1.4.2 The role of Mediator during transcription initiation and reinitiation

Mediator is recruited by activators to genes in a separable step from Pol II recruitment (Cosma et al., 2001), and it is thought to remain near the promoter, to facilitate transcription reinitiation (Liu et al., 2001; Yudkovsky et al., 2000). After the first transcription cycle, the scaffold complex, consisting of activator, Mediator and the general transcription factors TFIID, TFIIA, TFIIE and TFIIH, provides a platform for transcription reinitiation by Pol II. The scaffold complex enables genes to be transcribed more rapidly, because the slow step of PIC assembly is omitted. Transcription activation domains can stabilize the scaffold complex *in vitro* (Yudkovsky et al., 2000). An additional activation-independent function in initiation complex formation and transcription reinitiation has been attributed to the yeast Mediator tail module (Reeves and Hahn, 2003).

The Mediator CDK8/CyclinC module phosphorylates the CTD at serine 5 and stimulates the kinase activity of the general transcription factor TFIIH, which is the primary CTD kinase. CTD phosphorylation triggers the transition from transcription initiation to elongation (Kim et al., 1994; Li and Kornberg, 1994; Serizawa et al., 1993). In addition to CTD phosphorylation, TFIIH also phosphorylates the Mediator subunits MED4 and MED14 (Liu et al., 2004; Guidi et al., 2004), and CDK8/CyclinC phosphorylate Bdf1 and Taf2 (TFIID). The function of Mediator and TFIIH are linked to each other and both govern the CTD-dependent transcription in yeast extracts (Nair et al., 2005). Inhibition of the two kinases inhibits dissociation of the PICs (Liu et al., 2004), and therefore transcription initiation.

1.4.3 Structural information on Mediator complexes

Electron microscopy has been used to investigate the structure of Mediator alone and in complex with Pol II, activators or the CTD. Mediator unfolds in the presence of Pol II and adopts an extended conformation in which it wraps around the polymerase (Asturias et al., 1999). The whole polymerase is needed for Mediator unfolding since neither the polymerase without the CTD nor the CTD alone are able to induce complete unfolding. In this extended conformation, the biochemically defined head, middle and tail modules of Mediator can be distinguished. The fourth module, the CDK8/CyclinC module, was not present in these studies. In order to attribute the biochemically defined modules in the EM reconstructions, the wild type strain was compared to a

MED16 deletion mutant (Dotson et al., 2000; Davis et al., 2002). The whole tail module (MED2, MED3, MED15, MED16) is absent from Mediator in such a mutant. In addition, interaction maps obtained from yeast two hybrid analysis (Uetz et al., 2000; Guglielmi et al., 2004; Ito et al., 2001), coexpression in insect cells (Han et al., 2001; Kang et al., 2001), split ubiquitin assay (Gromoller and Lehming, 2000a) and urea washes (Han et al., 2001) gave indications about the Mediator subunit organization and allowed the attribution of the head and the middle module. No direct mapping of individual Mediator subunits has been done in yeast so far.

EM structural analysis of murine and human Mediator complexes revealed an overall structure similar to the yeast complex (Dotson et al., 2000). The CDK8/CyclinC module was not present in the structures of yeast and murine complexes, but is present in some structures of the human complex (Taatjes et al., 2002). There is no EM structure of a human Pol II/Mediator complex, but the effect of activators and the CTD on these complexes has been tested (Naar et al., 2002; Taatjes et al., 2002; Taatjes and Tjian, 2004). There are structural changes upon binding of each of the activators as well as upon binding of the CTD or the CDK8/CyclinC module. These conformational changes are relatively small compared to the unfolding that yeast Mediator undergoes upon binding to Pol II. The functional role of each of these conformational changes has not been investigated yet. They could represent a way to transfer signals from different activators to the polymerase (Chadick and Asturias, 2005). Furthermore, activators seem to stabilize the scaffold complex and thereby facilitate transcription reinitiation (Reeves and Hahn, 2003; Wang et al., 2005). It has been speculated that conformational changes have a role in PIC isomerization to allow promoter melting or promoter clearance (Struhl, 2005). In summary, even low resolution (~ 35 Å) electron microscopic studies show that Mediator undergoes strong structural changes upon interaction with Pol II and with transcription activators. The causes and consequences of these changes are however poorly understood, and the molecular mechanism of Mediator remained enigmatic, mainly because of a complete lack of detailed high resolution structural information.

1.5 The yeast Mediator proteins

1.5.1 The Mediator head module

The yeast Mediator head module consists of the proteins MED6, MED8, MED11, MED17 (Srb4), MED18 (Srb5), MED19 (Rox3), MED20 (Srb2) and MED22 (Srb6) (Boube et al., 2002). All four Srb genes were identified as dominant gain-of-function suppressors of a CTD truncation mutant (Thompson et al., 1993). While MED18 and MED20 are non-essential genes (Nonet and Young, 1989), MED17 and MED22 are essential (Thompson et al., 1993). Whole genome expression profiles for MED17 and MED22 show a general dependence of Pol II transcription on these genes (Holstege et al., 1998). MED18 and MED20 form a subcomplex, are required for the regulation of a specific subset of genes (van de Peppel et al., 2005), and seem to have a role in transcription initiation (Ranish et al., 1999). The other head module proteins (MED6, MED8, MED11, MED19) have been identified by peptide sequencing of purified complexes. MED6 is the highest conserved subunit of the head module (34 % homology between yeast and human). MED6, MED8, MED11 and MED19 are essential in *S. cerevisiae* and bind directly to the MED17 subunit, which plays a scaffold role within the module. *S. cerevisiae* MED8 is phosphorylated in the Mediator complex and as been shown to bind DNA at certain regulatory elements (Chaves et al., 1999). In addition to its occurrence in the Mediator, the mammalian MED8 can assemble with an elongin and a ringfinger protein (Cul2 and Rbx1, respectively) to reconstitute a Ubiquitin ligase elongin BC complex (Brower et al., 2002). Recombinant complexes of MED17/MED18/MED20/MED22 (Koh et al., 1998) and MED6/MED8/MED11/MED17/MED18/MED19/MED20/MED22 (Kang et al., 2001) have been obtained from baculo-virus expression systems. The latter binds to the Pol II CTD as well as to TBP and TFIIB, and it is able to stimulate basal transcription (two to three fold) but not activated transcription (Kang et al., 2001). The reconstituted head module can be associated with a reconstituted middle module to give a core Mediator complex. All head module proteins have homologous counterparts in higher eukaryotes.

1.5.2 The Mediator middle module

The Mediator middle module consists of the proteins MED1, MED4, MED7, MED9 (Cse2), MED10 (Nut2), MED21 (Srb7), MED31 (Soh1) and possibly MED5 (Nut1). MED4, MED7, MED9, MED10 and MED21 are essential in yeast while MED1, MED5, MED9 and MED31 are not essential (reviewed in Myers and Kornberg, 2000). The yeast Mediator middle module can be further divided into two biochemically stable submodules: the MED9 (MED1, MED4 MED9) and the MED10 (MED7, MED10, MED21) submodules (Kang et al., 2001; Han et al., 2001). Mutations of proteins in the two biochemical submodules have distinct phenotypes of transcriptional activation and repression (Han et al., 1999). In addition, Mediator purified from a MED9 deletion strain seems to be defective in stimulation of basal transcription (Liu et al., 2001). MED5 is the only protein of the middle module which has no homologous in higher eukaryotes, and is present only in a subpopulation of yeast Mediator. The MED5 protein has a histone acetyl transferase activity (Lorch et al., 2000). Its interactions with the middle module are not clear, even though yeast two hybrid screens indicate an interaction with MED1 in the middle and MED16 in the tail module (Guglielmi et al., 2004).

The MED31 protein was only recently discovered to be part of the Mediator middle module (Guglielmi et al., 2004; Linder and Gustafsson, 2004). Due to interactions with MED10 and MED21 observed by yeast two hybrid assays, MED31 most likely belongs to the MED10 submodule. It is note worthy that MED31 has 50 % sequence homology between yeast and human and is the best conserved core Mediator subunit. The MED31 protein is conserved also in *Giardia intestinalis* and *Plasmodium falciparum* which lack the CTD heptapeptide repeat (Linder and Gustafsson, 2004; Stiller and Hall, 2002). MED31 has therefore been suggested to be an ancient Mediator protein involved in signal transmission from regulators to polymerases in eukaryotes that diverged before the evolution of the CTD and the remaining Mediator subunits.

The other three proteins of the MED10 submodule, MED7, MED10, and MED21 are also highly conserved between yeast and human Mediator with 40, 39 and 44 % conservation, respectively. MED21 (Srb7) is the only protein of the core Mediator that has been discovered as a recessive mutation in a screen for mutants, which are able to restore viability of CTD truncations. Other recessive mutations found in this screen encode for proteins in the Mediator CDK8/CyclinC module.

Association of the Mediator middle module with the head module seem mediated through MED21 (Guglielmi et al., 2004; Gromoller and Lehming, 2000a). A reconstituted yeast Mediator middle module, containing all subunits but MED5 and MED31, was able to associate with the general transcription factors TBP, TFIIB and TFIIE as well as with the CTD (Kang et al., 2001).

1.5.3 The Mediator tail module

The Mediator tail module contains MED2, MED3 (Pgd1/Hrs1), MED14 (Rgr1), MED15 (Gal11), MED16 (Sin4). Of these subunits, only the N-terminus of MED14 is essential for viability. It binds to the MED9 submodule in the middle module and anchors the other subunits of the tail module to the core Mediator (Myers et al., 1999; Li et al., 1995).

Activators such as Gcn4 and Gal4 bind to subunits MED16 and MED15 (Myers et al., 1999; Bhoite et al., 2001; Han et al., 1999). Deletion of these subunits abolishes activated transcription by Gcn4, Gal4 or VP16 in yeast, but not basal transcription (Park et al., 2000). Recent studies in *S. cerevisiae* suggest that this module might function as a separate entity, which can be recruited to promoters independently of the core Mediator (Zhang et al., 2004; Bjorklund and Gustafsson, 2005). Additional activator independent functions of the yeast Mediator MED15 (Gal 11) module in PIC formation and reinitiation were found. Mutations in MED3 and MED16 destabilize PICs and lead to non-functional scaffold complexes (Reeves and Hahn, 2003). Mutations in MED3 lead to the partial disruption of the MED15 complex and therefore to defects in transcription reinitiation. In addition to its role in transcription activation, the Mediator tail module has a role in transcription repression. All proteins in the tail module (and some proteins in the head and middle module) have been discovered in screens for mutations affecting transcriptional regulation, mostly repression and not activation.

1.5.4 The CDK8/CyclinC module of Mediator

A subpopulation of Mediator complexes contain a CDK8/CyclinC module, which consists of CDK8/CyclinC, MED12 and MED13. This module is conserved in eukaryotes (Samuelson et al., 2003; Boube et al., 2002), and phosphorylates the CTD at serine 5 of the heptapeptide repeat (Borggreffe et al., 2002; Hengartner et al., 1998). *S. cerevisiae*

Mediator containing this module is implicated in transcription repression (Hengartner et al., 1998), *S. pombe* Mediator lacking this module has a stimulatory effect (Spahr et al., 2003). The holoenzyme complex consisting of Pol II and the Mediator does not contain the CDK8/CyclinC module (Samuelson et al., 2003). The human Mediator complexes NAT and SMCC containing the CDK8/CyclinC module are not capable of activated transcription (Gu et al., 1999; Sun et al., 1998). There are also reports on positive effects of the CDK8/CyclinC module on transcription by the recruitment of Gal4 activation domain (Ansari et al., 2002) and the phosphorylation of the gene-specific activator Sip4 (Vincent et al., 2001). CDK8/CyclinC phosphorylates TFIID subunits (Liu et al., 2004) as well as MED2 (Hallberg et al., 2004). Mutation of the MED2 phosphorylation site reduces expression levels of certain genes. The CDK8/CyclinC module promotes ATP dependent dissociation of preinitiation complexes to the scaffold complex (Liu et al., 2004). In addition to its function in transcription, the CDK8/CyclinC complex plays also a role in cell cycle regulation in human cells. (Ren and Rollins, 2004; Sage, 2004; Akoulitchev et al., 2000). The structure of CyclinC has recently been solved (Hoepfner et al., 2005) and CDK8 was modeled from the existing structure of CDK7 (Hoepfner et al., 2005).

1.5.5 MED25, an additional human Mediator subunit

Mediator complexes of higher eukaryotes contain several additional proteins compared to the yeast Mediator (MED23 to MED30) (Bourbon et al., 2004, Table 2). MED25 (Arc92/ACID1) was identified as a target of the Herpes simplex virus transactivator VP16 (Mittler et al., 2003; Yang et al., 2004) in human and it has been copurified with Mediator (Naar et al., 1999). The MED25 protein is conserved in mammals, flies, fish and chordata and is present in only a subpopulation of human Mediator complexes. Gst pull-down assays identified Dif and HSF as cellular MED25 interacting activators in *Drosophila* (Kim et al., 2004; Kim and Lis, 2005). MED25 associates with the Mediator through a von Willebrand type A domain which is linked through a proline- and glutamine-rich linker to its activator interaction domain (ACID). Prostate tumor overexpressed gene (PTOV-1) is the only gene in the database with sequence homology to ACID. PTOV-1 contains two ACID copies.

1.6 Herpes Simplex Virion Protein 16 (VP16), a viral activator of transcription

Transcriptional activator proteins regulate gene expression in response to environmental changes. Activator proteins have a modular architecture. They generally contain two domains: a DNA-binding domain which recognizes specific sequences and an activation domain which recruits components of the transcription machinery and chromatin modifying factors. Because activation domains can act independently of the DNA-binding domains, artificial chimeras containing a heterologous DNA-binding domain are often used (Sadowski et al., 1988). One example is the Gal4-VP16 fusion protein, where a Gal4 DNA-binding domain is fused to the VP16 activation domain. VP16 contains an activation domain at the C-terminus (residues 411-490), which can be subdivided into two regions, H1 (residues 411-452) and H2 (residues 453-490). H1 and H2 can activate transcription independently *in vivo*, and they act synergistically (Walker et al., 1993). The VP16 activation domain belongs to the group of acidic activation domains. Initially, acidic activation domains were thought to activate transcription as “acidic blobs”, which contain an amphipathic α -helix with an acidic face (Sigler, 1988). Later, however, additional motifs were discovered to be critical for activation functions. In VP16, hydrophobic phenylalanines are important for activation (Cress and Triezenberg, 1991; Sullivan et al., 1998).

The H2 region of the activation domain undergoes an induced transition from random coil to α -helix upon binding to a human TAF (TBP associated factor) (Uesugi et al., 1997). Formation of an α -helix in the VP16 activation domain is also observed upon binding to PC4 (Jonker et al., 2005). While α -helix formation was clearly visible for the C-terminal H2 part of the activation domain, the N-terminal H1 part showed hardly any secondary structure in this study. Among several other targets within the transcriptional machinery and chromatin-associated factors, VP16 binds to the mammalian Mediator complex *via* MED25 (Mittler et al., 2003; Yang et al., 2004). VP16 recruits MED25 and Mediator and thereby activates transcription. This interaction seems to be important *in vivo*, since the ACID domain of MED25 acts in a dominant negative way (Mittler et al., 2003; Yang et al., 2004).

1.7 Aims and scope of this work

High-resolution structural information of Mediator is required in order to gain insights into the molecular mechanism of this fundamental complex. To obtain X-ray structures, large amounts of high quality protein samples have to be purified. The modular structure as well as the low abundance of Mediator within the cell have so far hampered X-ray structure determination of the natively purified complex. Single Mediator subunits cannot be easily overexpressed in recombinant *E. coli* systems, mainly due to the hydrophobic nature of proteins occurring in such large complexes.

The aim of the project was to establish a recombinant *E. coli* expression system to analyze pairwise interactions between proteins belonging to the Mediator middle module, in order to generate subcomplexes suitable for crystallization. A short-term objective was the crystallization and subsequent structure determination of such subcomplexes. A longer-term objective was the reconstitution of a recombinant Mediator middle module. Such a recombinant complex could then be used for biochemical characterization and to solve higher resolution structures of the Mediator middle module with and without Pol II. Identification of the subunits contacting the Pol II is an important step in understanding how Mediator regulates the polymerase. A second aim of the work was to take first steps towards understanding how activators contact the Mediator. A previously identified activator interaction domain of the mammalian Mediator subunit Med25 (Mittler et al., 2003; Yang et al., 2004) was analyzed for its binding to the viral activator VP16.

2 Results

2.1 The MED7/MED21 subcomplex: Structure and protein interactions

2.1.1 Subcomplex domain mapping

Individually expressed recombinant Mediator subunits are generally insoluble, explaining the current lack of Mediator subunit structures. Insolubility apparently results from a loss of structural integrity when subunits are outside their natural multi-protein context. To overcome this obstacle, I coexpressed the highly conserved and essential *S. cerevisiae* Mediator subunits MED7 and MED21 in *E. coli* with the use of a bicistronic vector (Experimental procedures 4.2). Purification by a His tag on Med21 copurified MED7. To probe for flexible regions that may prevent crystallization the purified MED7/MED21 complex was subjected to partial proteolysis. Chymotrypsin treatment resulted in the removal of 101 poorly conserved N-terminal residues of MED7 (Figure 2A). A corresponding variant MED7 Δ N (MED7 residues 102-222) still formed a stable complex with MED21 after coexpression. Sequence alignment showed a non-conserved C-terminus of MED7 after amino acid 205. A subsequently prepared variant MED7 Δ N Δ C (residues 102-205) still bound MED21 strongly. This variant comprises only the highly conserved region of MED7, which shows 59 % sequence homology between the yeast and human proteins. Except for a short C-terminal truncation, MED21 remained stable in all proteolysis experiments (Figure 2B). In contrast, individually expressed MED21 was readily cleaved before residues 38 and 76. Therefore, MED21 is protected from degradation upon MED7 binding. Taken together, iterative proteolysis and truncation of coexpressed and copurified subunits allowed us to map a stable subcomplex, and this approach may be used to obtain potentially crystallizable portions of other multi-protein complexes.

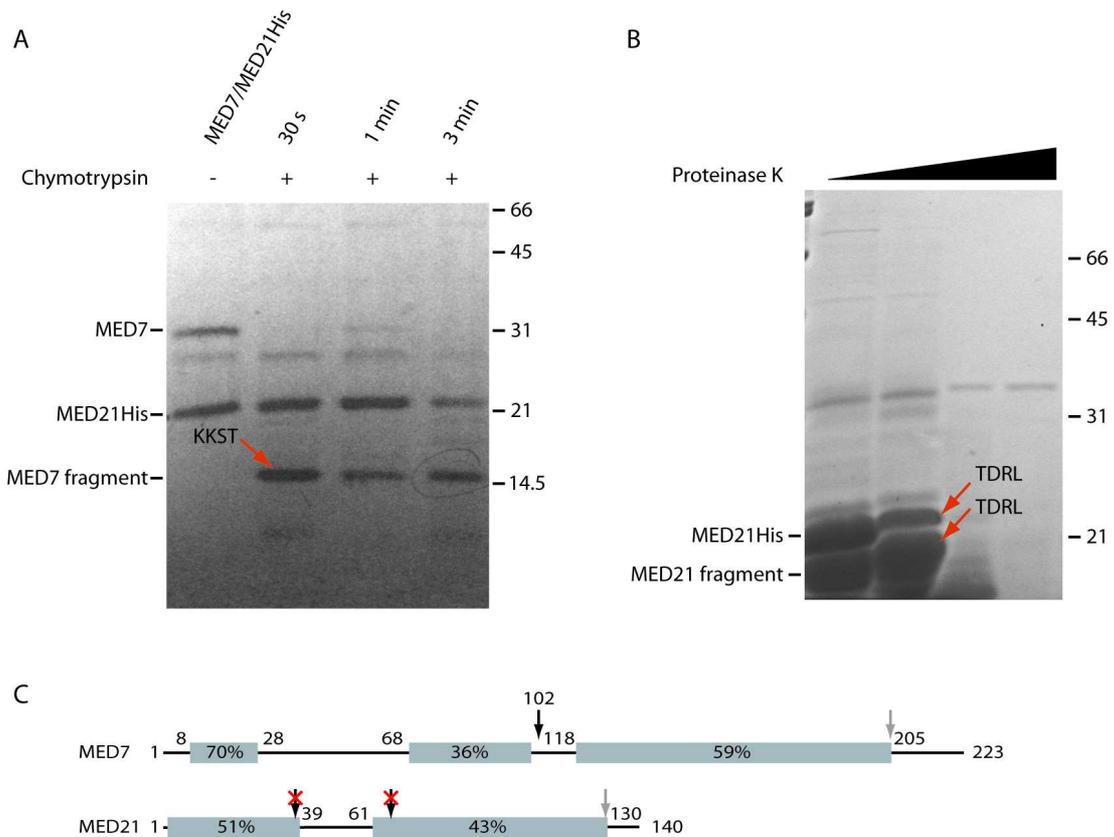
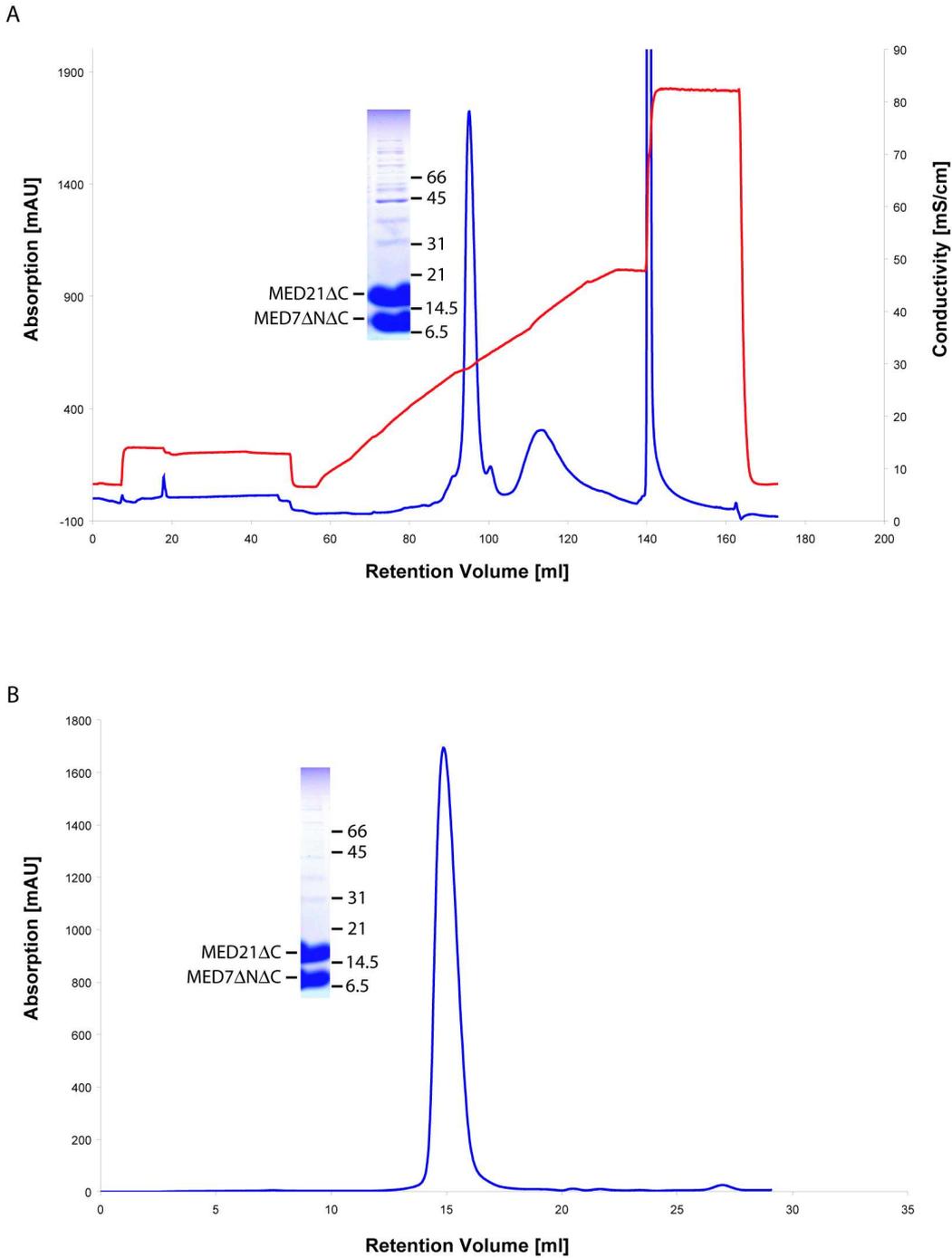


Figure 2: Partial proteolysis of MED7/MED21 complex: A) Coomassie-stained SDS-PAGE of a chymotrypsin digest of MED7/MED21 complex. Arrows mark bands analyzed by Edman sequencing. The resulting N-terminal sequence is given in one letter code. B) Coomassie-stained SDS-PAGE of a proteinase K digest of MED7 Δ N Δ C/MED21. The sequenced bands are indicated by arrows. C) Schematic diagram of the MED7 and MED21 proteins. Blue bars indicate regions of homologous sequences in different species. Numbers correspond to the percentage of homology between yeast and human. Proteolytic cleavage sites are indicated above the schemes with black (sequenced) or grey arrows (estimated from SDS-PAGE). Red crosses indicate the protection of the cleavage site by the interaction partner.

2.1.2 Purification of crystallizable MED7/MED21 variants

The exact protocols for purification of the different variants of the MED7/MED21 complex are described in Experimental Procedures (4.1.5). As the protein complex tends to form aggregates at a pH below 7.5, all steps throughout the purification were carried out at pH 8.5. First protein was enriched either by Ni-NTA affinity purification (MED7 Δ N Δ C/MED21) or by ammonium sulfate precipitation (MED7 Δ N Δ C/MED21 Δ C). Ammonium sulfate precipitation is a very powerful purification step since the proteins of the Mediator middle module precipitate below 35 % saturated ammonium sulfate, a concentration in which most *E.coli* proteins are still soluble. The pellet can be collected, proteins redissolved and further purified by anion exchange (Figure 3A) and gelfiltration chromatography (Figure 3B). Anion exchange chromatography is important to separate the MED7/MED21 complex from excess MED21, which does not bind the column. The gelfiltration step is not needed for purification of the complex but is used to remove eventual aggregates and ensure complete buffer exchange to the buffer used for crystallization. The quality of the protein preparation is very high (Figure 3B) and the complex is easy to handle. Aliquots of 50 μ l could be flash frozen in liquid nitrogen and stored at -80° C for later use.

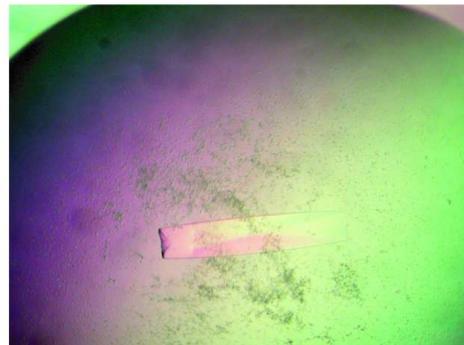


2.1.3 Crystallization of the MED7/MED21 heterodimer

Preparations of the MED7 Δ N Δ C/MED21 subcomplex were monodisperse and homogeneous according to dynamic light scattering and size exclusion chromatography. Crystallization trials using commercial Hampton screens were set up with the hanging drop vapor diffusion method. Microcrystals were obtained and could be refined by varying salt and ethanol concentrations as well as screening various additives. These screens lead to the observation that addition of β -mercaptoethanol improves crystal size and quality. The final crystallization condition was: 0.8-1.1 M NaCl, 10-15 % ethanol, 0-3 % PEG 6000, 10 mM β -mercaptoethanol, 10 mM EDTA. Crystals grew to a maximum size of 0.6 x 0.3 x 0.3 mm and comprise two heterodimers per asymmetric unit (space group C222₁) (Figure 4). The crystals were very sensitive to cryo-cooling and a slow cryo-protocol using microdialysis buttons had to be established. Despite the large size of the crystals, diffraction extended only to 3.3 Å in favorable cases.



Hampton screen 2 #8



refined condition

Figure 4: Crystals of the MED7 Δ N Δ C/MED21His complex obtained in the Hampton screen 2 condition 8 (1.5M NaCl, 10% ethanol, left) and in the refined condition (0.8 M NaCl, 10 % ethanol, 3 % PEG 6000, 10 mM β -mercaptoethanol, 10mM EDTA, right). This crystal grew to a maximal size of 50x 50x 150 μ M.

To increase the resolution limit of the crystals, I truncated eight non-conserved amino acid residues from the MED21 C-terminus (variant MED21 Δ C, residues 1-132), which were unstable in protease digest experiments (Figure 2B). The resulting complex MED7 Δ N Δ C/MED21 Δ C crystallized in additional conditions, not containing ethanol

(see Experimental procedures 4.2.5, Figure 5). Crystals adopted space group $P4_322$. These crystals were slightly more stable, and could be used for structure determination.

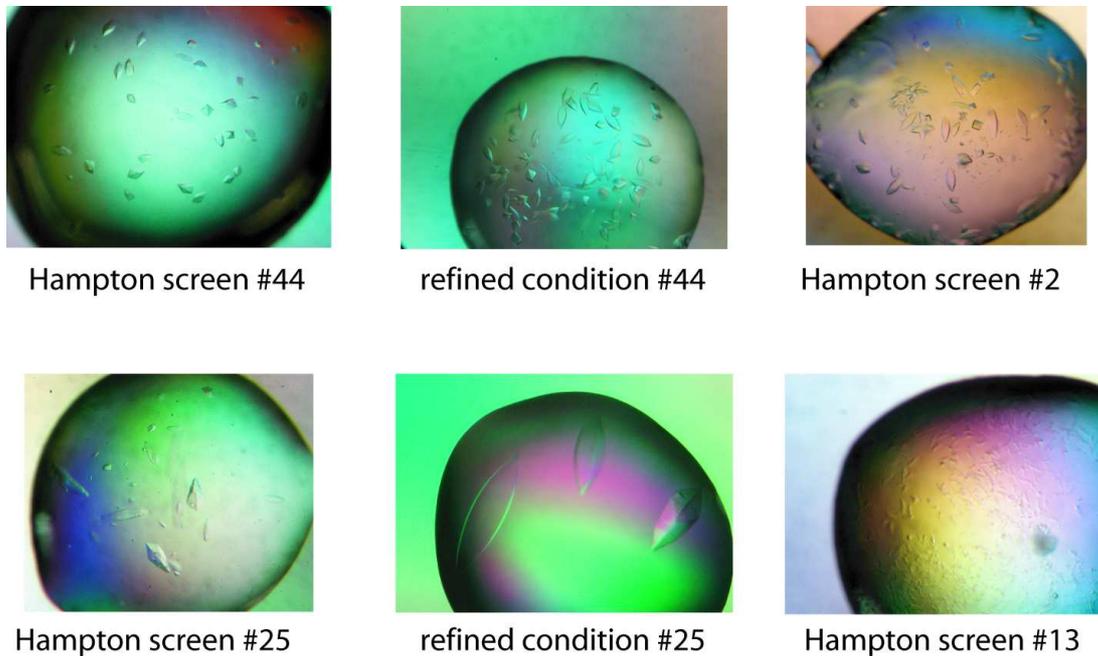


Figure 5: Crystals of the MED7 Δ N Δ C/MED21 Δ C complex obtained in the Hampton screen and in the refined condition (Experimental Procedures 12).

2.1.4 Methionine mutations enhance the anomalous signal

Initial attempts of structure determination by multiple anomalous dispersion (MAD) were done using selenomethionine substituted crystals of MED7 Δ N Δ C/MED21His. Data to 3.8 Å resolution were obtained. The anomalous signal of the crystal containing seven methionines per heterodimer (including the N-termini) was strong enough to find three selenium peaks and obtain initial phases. However the quality of these phases was very poor and the resulting electron density map could not be used for model building. To increase the anomalous signal and thus obtain better phasing power I replaced three leucines in MED21 by methionines (L5M, L119M, L125M). The positions for the mutations were chosen such that they lay in conserved regions with predicted secondary structure and could serve as sequence markers during model building. MAD data to 3.6 Å resolution was obtained from the resulting crystals and could be used for phasing (Table 3).

Crystal	SeMet MAD			Native	Native
	MED7 Δ N Δ C/ MED21 Δ C ¹			MED7 Δ N Δ C/ MED21 Δ C ¹	MED7 Δ N Δ C/ MED21
Data collection					
Space group	P4 ₃ 22			P4 ₃ 22	C222 ₁
a (Å)	85.6			85.7	121.5
b (Å)	85.6			85.7	128.9
c (Å)	185.2			183.0	170.2
Wavelength (Å)	0.9795 peak	0.979	0.9686	0.9795	1.0749
		inflection	remote		
Resolution range	20-3.6 (3.73- 3.6) ²	20-3.6 (3.73- 3.6)	20-3.6 (3.73- 3.6)	50-3.0 (3.11- 3.0)	50-3.3 (3.42- 3.3)
Completeness (%)	100 (100)	100 (100)	100 (100)	98.6 (100)	99.6 (99.5)
Unique reflections	8517	8517	8517	14180	20641
Redundancy	6.9	6.5	6.7	5.4	4.8
R _{sym} (%)	8.3 (21.0)	8.7 (32.0)	7.8 (26.0)	6.5 (26.2)	5.8 (36.2)
<I/σ>	7.6	6.4	6.8	8.9	10.4
f' ³	-8.0	-9.8	-3.8		
f'' ³	4.9	2.5	3.7		

Table 3: Diffraction data

¹The triple point mutant L5M/L119M/L125M of MED21 was used.

²The numbers in parenthesis correspond to the highest resolution shell.

³The values for f' and f'' were obtained from a fluorescence scan of the derivative crystals.

2.1.5 Structure determination of the MED7/MED21 heterodimer

The structure was solved with MAD data from a selenomethionine-substituted crystal of MED7 Δ N Δ C/MED21 Δ C containing the mutations, and was refined to a free R-factor of 28.9 % against native diffraction data extending to 3.0 Å resolution (Experimental procedures, Table 4). I used the refined structure of the MED7 Δ N Δ C/MED21 Δ C heterodimer to solve the structure in the initial crystal form (space group C222₁, MED7 Δ N Δ C/MED21) by molecular replacement. I refined this structure to a free R-factor of 31.4 % with data extending to 3.3 Å resolution (Table 4).

Crystal	Native MED7 Δ N Δ C/MED21 Δ C ¹	Native MED7 Δ N Δ C/MED21
Number of residues	209	428
Non-hydrogen atoms	1737	3549
RMSD bonds (Å)	0.008	0.009
RMSD angles (°)	1.164	1.37
Ramachandran plot ² (core/allowed/additionally allowed)	94.3/5.7/0	87.9/11.8/0.3
R _{cryst} (%)	25.7	27.9
R _{free} (%)	28.9	31.4

Table 4: Refinement statistics

¹The triple point mutant L5M/L119M/L125M of MED21 was used.

²PROCHECK (56).

2.1.6 The MED7/MED21 structure

Elongated two domain structure

MED7 and MED21 form a very elongated heterodimer of purely helical structure (Figure 6). The heterodimer extends over 110 Å, corresponding to one third of the Mediator length (Davis et al., 2002), and amounting almost to the diameter of Pol II (Cramer et al., 2001) (Figure 6). This elongated shape is consistent with an unusually short retention in size exclusion chromatography (Figure 7). The retention of the MED7/MED21 species lies between that of the standard proteins aldolase (157 kDa) and catalase (230 kDa), and may be explained by formation of tetramers (dimers of heterodimers). Two different types of tetramers are observed in the crystal, which show a computed (Garcia De La Torre et al., 2000) Stokes radius of 42-44 Å. This is in good agreement with the Stokes radii for aldolase and catalase (47 Å and 49 Å, respectively), whereas the Stokes radius of the MED7/MED21 heterodimer is only 31 Å. Thus the complex apparently forms stable tetramers in solution. Deletion of the C-terminus of MED21 (MED21 1-102) results in longer retention on size exclusion chromatography, indicating that the C-termini are responsible for dimerization. The C-termini are also involved in the formation of one of the heterotetramers in the crystals. Static light scattering also indicates a molecular weight of about 60kDa and a hydrodynamic radius of

49 Å, consistent with a heterotetramer. MED7 and MED21 show no sequence similarity, but they have the same structural organization, comprising three extended helices. MED7 and MED21 tightly pack against each other, forming a heterodimer with two domains. A 4-helix bundle domain is formed by the two N-terminal α -helices of each subunit, and the C-terminal α -helices of the two subunits form a long coiled-coil protrusion (Figure 6). A “hinge” region connects the bundle domain to the coiled-coil protrusion.

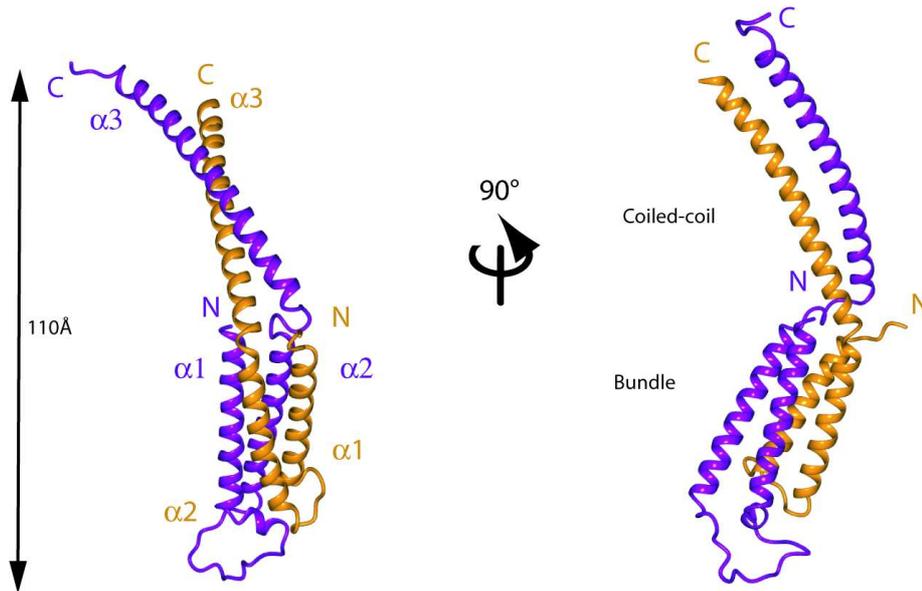


Figure 6: MED7/MED21 (Med7/Srb7) structure: Two views of a ribbon model of the MED7/MED21 complex, related by a 90° rotation around the vertical axis. MED7 is in orange, MED21 in purple. Figure prepared with DINO (<http://www.dino3d.org>).

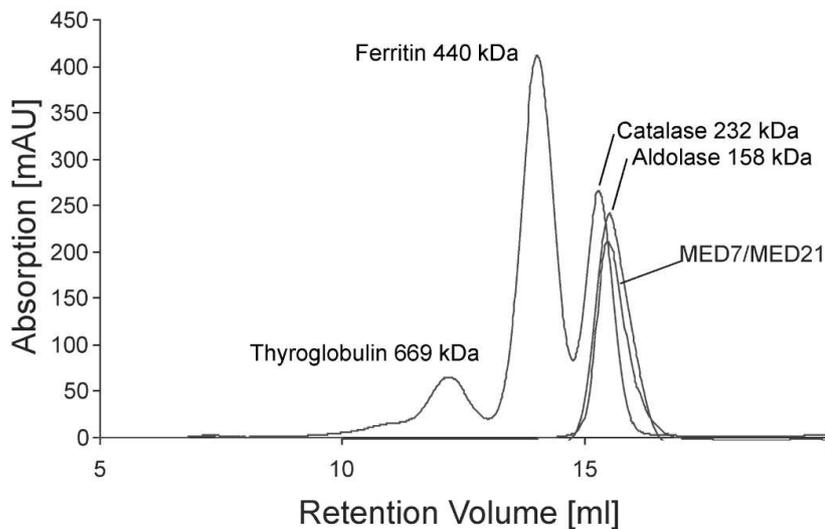


Figure 7: Size exclusion chromatography: The MED7/MED21 complex shows an unusually short retention in size exclusion chromatography. The elution profile of MED7/MED21 on a Superose-6 column (Amersham) is overlaid to those of standard proteins (thyroglobulin, ferritin, catalase, and aldolase). Protein elution was monitored by UV absorption at 280 nm. For details see Experimental procedures.

A conserved flexible hinge Comparison of the MED7/MED21 structure in the two crystal forms reveals that the coiled-coil protrusion can undergo a hinge movement with respect to the bundle domain, giving rise to a 10 Å-displacement of the C-terminal end of the protrusion (Figure 8). This repositioning of the protrusion is accommodated by a slight conformational adjustment in the hinge region. Thus the MED7/MED21 heterodimer has an intrinsic flexibility that allows for a relative repositioning of the two domains. This flexibility is not due to the short C-terminal truncation of MED21 in the C222₁ crystal form, since the two chemically identical heterodimers in the asymmetric unit also show two different conformations indicating the same hinge motion, and since the truncated residues are not involved in packing interactions in the C222₁ crystal form. The high sequence conservation of amino acid residues in the hinge region of MED7 and MED21 (Figure 9A) strongly suggests that the observed flexibility is functionally significant. A hinge movement is also predicted by molecular dynamics simulation with the Dynamite server (<http://dynamite.biop.ox.ac.uk/dynamite>).

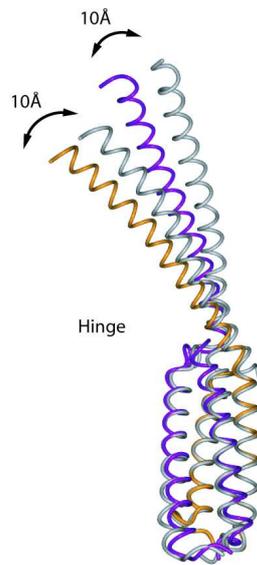


Figure 8: MED7/MED21 hinge. Structures in two different crystal forms have been superimposed with their bundle domains, resulting in a 10 Å difference at the end of the coiled-coil in the $C222_1$ crystal form (color) compared to the $P4_322$ crystal form (grey).

MED7-MED21 interaction MED7 and MED21 form a continuous hydrophobic interface, which explains why subunits that are individually expressed in *E. coli* are structurally unstable and either poorly soluble or easily degraded. The hydrophobic residues in the MED7-MED21 interface along the coiled-coil protrusion are mostly conserved throughout eukaryotes and show a spacing typical for coiled coils (MED7 residues L175, L178, L182, I189, I192; MED21 residues I94, L97, L101, V104, A111) (Figures 9A, B). In MED7, the regular spacing of hydrophobic residues is discontinued only at residue K185. The regular pattern in MED21 is similar, except residues Q90, K108, and K115, which however also participate in fold-stabilizing interactions. The conserved MED21 residue Q90 forms hydrogen bonds with E172 and R171 of MED7, and may contribute to the specificity of the heterodimeric interaction. Given the high degree of sequence conservation of the crystallized regions, the MED7/MED21 structure must essentially be the same in all species. This conservation in structure explains why a chimeric MED21 protein consisting of the two human N-terminal helices (corresponding to yeast residues 1-81) and the yeast C-terminal helix (residues 82-140) is functional *in vivo* (Chao et al., 1996), and why truncation of the MED21-interacting region of MED7 is lethal in *C. elegans* (Kwon et al., 2001).

A unique leucine zipper The coiled-coil protrusion resembles a canonical leucine zipper as it occurs in transcription factors like c-Jun (Junius et al., 1996) or Gcn4 (Keller et al., 1995; Holm and Sander, 1995). The two helices of the coiled-coil separate at their C-terminal end. The inner side of this open end of the zipper is conserved and hydrophobic (MED21 residues L118, V122, I126, and F129; MED7 residues L203 and V199), and thus chemically different from the corresponding region in leucine zipper-containing transcription factors, which show exposed polar or charged residues for DNA binding in this region. Instead of binding DNA, the open end of the MED7/MED21 coiled-coil may interact with other Mediator subunits. In both crystal forms, two heterodimers pack against each other via their coiled-coil ends (FIG. 9C), indicating that the open end of the coiled-coil may allow for protein interactions.

2.1.7 Conserved interactions of the MED7/MED21 heterodimer within the Mediator middle module

To map direct protein-protein interactions of the MED7/MED21 heterodimer with other subunits in the Mediator middle module, I tested for copurification of subunits after their coexpression in *E. coli* (Figure 10A). Such copurification can successfully map strong and specific direct protein-protein interactions, as demonstrated in the structure determination of the MED7/MED21 complex. The copurification assay is very stringent, since many different non-specific competitor proteins are present in the *E. coli* lysate, since the stoichiometry of the complexes can be estimated with Coomassie-stained gels, and since the protein-protein complexes must persist over several copurification steps, even when elevated salt concentrations of 600 mM NaCl are used. *E. coli* was cotransformed with two plasmids, a bicistronic plasmid expressing the MED7/MED21 heterodimer, and a second plasmid with different antibiotic resistance, expressing a third subunit. I found that the MED7/MED21 heterodimer strongly binds MED10 and MED4, which show 39 % and 33 % sequence homology between yeast and human, respectively, and belong to the most conserved core Mediator subunits together with MED7, MED21, MED31 (Soh1) and MED6 (Figure 10A). In contrast, the other two subunits of the middle module, MED1 or MED9 do not copurify with the MED7/MED21 complex. To test if MED21 alone is sufficient for the interactions with MED10 or MED4, I constructed bicistronic vectors for coexpression. These

experiments revealed that MED21 alone is capable to bind MED10 or MED4 (Figure 10B).

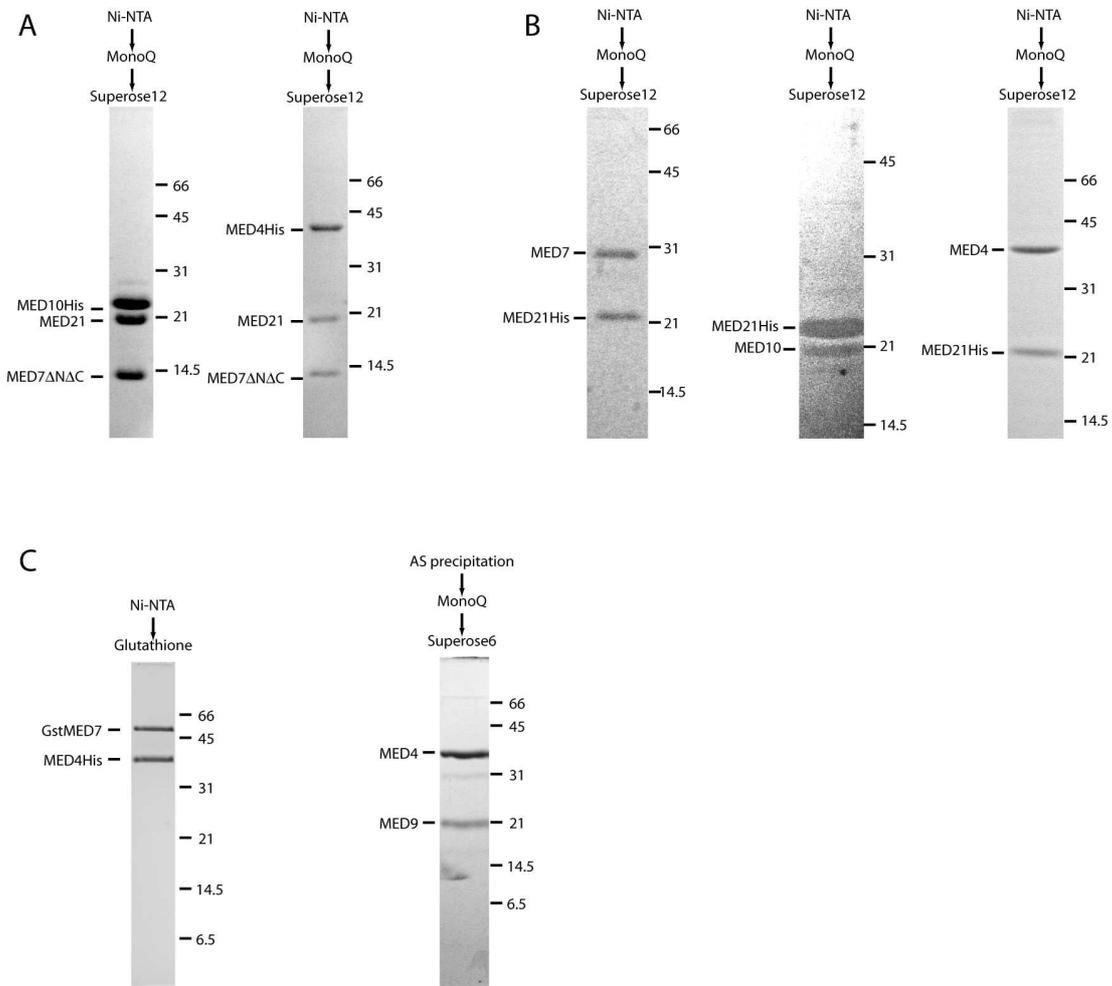


Figure 10: Protein-protein interactions of the MED7/MED21 complex. (A) Copurifications of MED7ΔNΔC/MED21 heterodimer with either MED10His or MED4His. (B) Copurification of MED21His with MED7, MED10 and MED4. (C) Copurifications of MED4His with GstMED7 and MED9. Gels were stained with Coomassie Blue.

2.1.8 MED4 and MED9

Urea dissociation experiments of yeast Mediator middle module subunits by Han et al. (Han et al., 2001) identified two stable subcomplexes within the Mediator middle module. One consisting of MED7/MED10/MED21 was already reconstituted earlier

in this study. The three other subunits of the middle module, MED1, MED4 and MED9 also form a stable subcomplex. Coexpression of MED4His and MED9 leads to massive overexpression of both proteins. Subsequent purification of MED4 by His affinity chromatography copurifies MED9 (Figure 10C). Coexpression of a plasmid encoding for the genes of MED4 and MED9 with a plasmid encoding for MED1His did not allow for the copurification of MED4/MED9. MED4 also interacts with MED7. Purification of MED4 and MED7 using a His tag on MED4 leads to the copurification of both proteins. In a second purification step using an N-terminal Gst tag on MED7 a stoichiometric MED4/MED7 complex is purified (Figure 10C). This copurification is not due to the Gst-tag since MED4 also copurifies MED7 if Gst is not present. It is interesting to note that the expression of GstMED7 is only possible in the presence of MED4. Together with the interaction of MED4 and MED21 shown in 2.1.7 this establishes the interaction between the two subcomplexes of the middle module.

2.1.9 MED6 bridges two Mediator modules

I could additionally demonstrate that the MED7/MED21 heterodimer binds directly to MED6, which shows 34 % sequence homology between yeast and human (Figure 11A). While MED7, MED21, MED4 and MED10 are all subunits of the middle module, MED6 is an integral part of the head module (Lee and Kim, 1998), suggesting that MED6 bridges these two modules. To test if MED6 binds directly to MED17 (Srb4), the architectural subunit of the head module (Koh et al., 1998), I tagged MED6 with a C-terminal hexahistidine tag (His), fused MED17 to a N-terminal Gst tag (Gst), and coexpressed the two subunits from a bicistronic vector. I could copurify the two subunits in two subsequent affinity chromatography steps, using a Ni-NTA and a glutathione column (Figure 11A, B). Successful purification of the complex was independent of the order of the affinity columns. The weakly conserved N-terminal part of MED17, and the non-conserved C-terminal part of MED6, are not required for binding since truncated variants of MED17 (residues 241-688) and MED6 (residues 1-214) were sufficient for the interaction. These results are consistent with a functional interaction between MED17 and MED6 observed previously (Lee and Kim, 1998). The results are further consistent with a very recent study of Mediator subunit interactions by yeast two-hybrid analysis (Guglielmi et al., 2004). In conclusion, MED6 physically bridges between the two Mediator core modules, interacting with MED17 in the head module and with the MED7/MED21 heterodimer in the middle module.

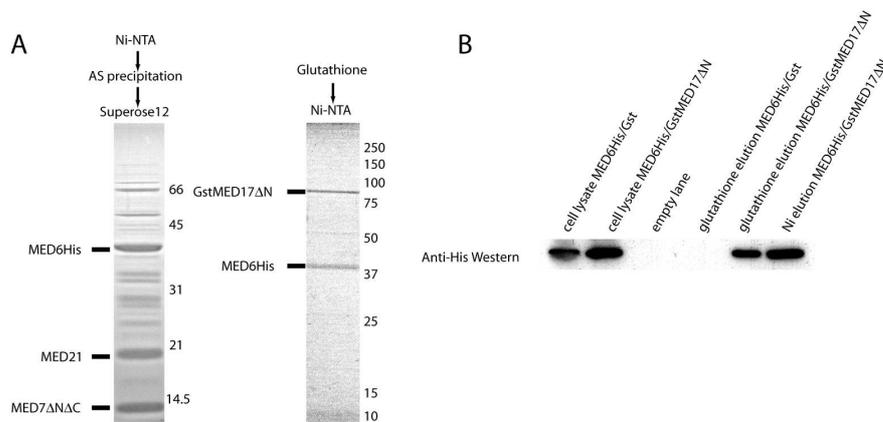


Figure 11: (A) Copurification of MED6His with MED7 Δ N Δ C/MED21 or GstMED17 Δ N. A schematic presentation of the purification procedure is shown above the gels. Gels were stained with Coomassie Blue. (B) Western blot of the GstMED17 Δ N/MED6His purification. The binding of MED6 to MED17 is not due to the presence of the Gst tag since a purification using Gst only does not yield MED6 (fourth lane).

2.1.10 MED31: an integral part of the middle module

While this study was underway, an additional yeast Mediator subunit, MED31 (Soh1) was identified (Guglielmi et al., 2004; Linder and Gustafsson, 2004). MED31 has been known to be part of murine Mediator complexes. It has very strong homology of 50 % between yeast and human which makes it to the most conserved Mediator subunit except for CDK8/CyclinC (Linder and Gustafsson, 2004). Yet no purification of yeast Mediator has found MED31 associated to it before.

According to yeast two hybrid screens performed by Guglielmi et al. this additional subunit interacts with MED10 and MED21 and it therefore seems to be an integral part of the middle module (Guglielmi et al., 2004). To test for direct protein protein interactions, I coexpressed MED31 together with subunits of the middle module and MED6 in *E. coli* and pulled on the proteins by a His tag on one of the subunits (Figure 12A). The results indicate an interaction of MED31 with either MED4, MED7, MED10 or MED21, but not with MED6. As a negative control overexpressed MED31 was incubated with Ni-NTA beads. I could therefore confirm that MED31 is an integral part of the middle module in yeast Mediator.

Coexpression and purification of MED10His or MED7 Δ N Δ C/MED21His with MED31

leads to copurification of MED31. However, the complexes are not stable during purification and tend to aggregate. Limited proteolysis of purified MED10His/MED31 using chymotrypsin, trypsin or proteinase K reveals that MED31 is very stable (Figure 12B). MED10 in contrast is degraded after 1 minute of incubation with any of the proteases used. Having found a very stable Mediator protein I expressed, purified and crystallized MED31 alone. MED31 can be expressed and purified in very small amounts. Crystals could be obtained in initial crystallization setups in condition 7 of the Nextal classic screen (100 mM Na citrate pH 5.6, 20 % isopropanol, 20 % PEG 4000). These crystal grew at a maximal size of about 20 x 20 x 10 μm . However, they did not diffract on a home source and were not easily reproducible.

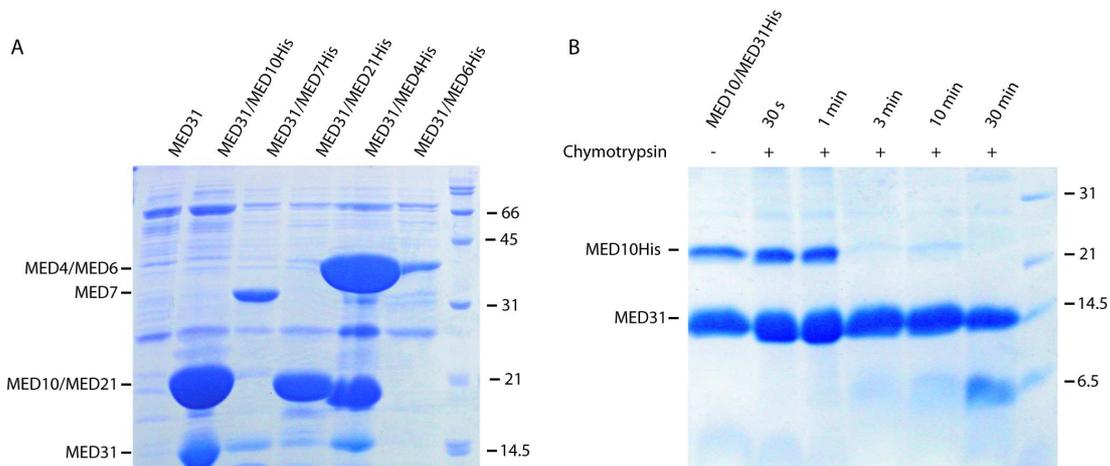


Figure 12: MED31 interacts with several subunits of the middle module: (A) Pull down of MED31 with MED4, MED6, MED7, MED10 and MED21 with the use of a His tag on these proteins. The names of the Proteins are indicated on the left, the molecular weight (kDa) on the right side of the gel. (B) Limited proteolysis of copurified MED10/MED31 complex by chymotrypsin. Proteins and the molecular weight (kDa) are indicated. Polyacrylamide gels are stained with Coomassie blue.

2.1.11 Reconstitution of larger subcomplexes of the Mediator middle module

Given the central role of the MED7 Δ N Δ C/MED21 complex within the Mediator middle module it could serve as a scaffold for crystallization and structure determination of larger complexes. Taking advantage of the interaction partners determined by the copurification assay, I reconstituted trimeric subunit complexes containing the MED7 Δ N Δ C/MED21 heterodimer together with either MED4 or MED10 as well as a tetrameric complex containing MED4 and MED9.

Complexes of MED7 Δ N Δ C/MED10/MED21 or MED4/MED7 Δ N Δ C/MED21 were purified with the use of the His tags on MED4 or MED10. Nickel affinity and anion exchange chromatography resulted in stoichiometric MED7 Δ N Δ C/MED10/MED21 complex, which elutes as one peak from a size exclusion column.

MED4/MED7 Δ N Δ C/MED21 was not purified as a stoichiometric complex by Ni affinity and anion exchange chromatography. An excess of monomeric MED4 was removed from the complex by two subsequent rounds of size exclusion chromatography. A complex of MED4/MED7 Δ N Δ C/MED9/MED21 was purified in the same way. All purified complexes were used for crystallization trials.

Analytical size exclusion chromatography shows an unexpected short retention volume of the trimeric complexes, similar to that of the MED7/MED21 heterodimer (Figure 13). MED10 does not contribute much to the overall shape and its retention volume is only slightly larger than that of MED7 Δ N Δ C/MED21, indicating that it may form a stable heterohexamer (dimer of trimers). The complex containing MED4 has a retention time which lies between that of the standard proteins ferritin (440 kDa) and catalase (232 kDa). A complex containing MED4 seems to have an even more pronounced elongated shape than the MED7 Δ N Δ C/MED21 heterodimer. Inclusion of MED9 into a complex containing MED4/MED7 Δ N Δ C/MED9/MED21 leads to a retention time slightly shorter than that of ferritin. This corresponds to a even more elongated shape. Taken together this indicates that the dimerization seen in the MED7 Δ N Δ C/MED21 complex persists also in complexes containing additional subunits. Alternatively, that one of the heterodimers is replaced by the new subunits which still have very elongated shapes.

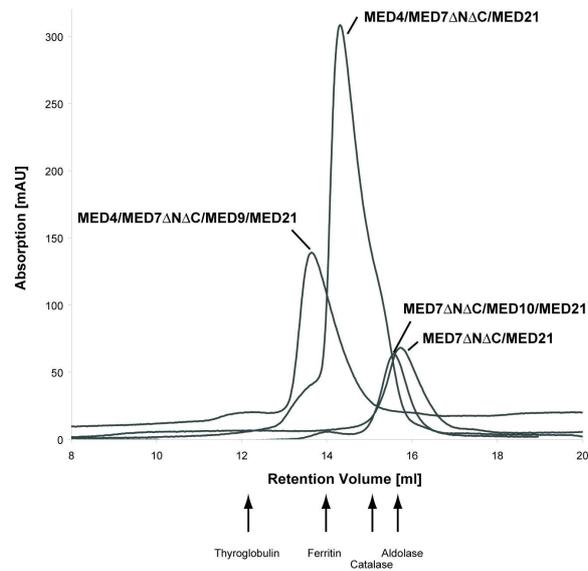


Figure 13: Size exclusion chromatography of subcomplexes of the Mediator middle module: The complexes containing MED7/MED21 and MED4, MED4/MED9 or MED10 show unusually short retention in size exclusion chromatography. The elution profiles of the complexes on a Superose 6 column (Amersham) are overlaid. The elution of the standard proteins thyroglobulin (660 kDa), ferritin (440 kDa), catalase (232 kDa), and aldolase 158 kDa) is indicated by arrows below the chromatogram. Protein elution was monitored by UV absorption at 280 nm. For details see Experimental procedures.

Crystallization of MED7 Δ N Δ C/MED10/MED21 and MED4/MED7 Δ N Δ C/MED21

The MED7 Δ N Δ C/MED10/MED21 complex crystallized under the same conditions as the MED7 Δ N Δ C/MED21 complex using the hanging drop vapor diffusion method. The presence of MED10 was confirmed by mass spectrometry of washed and dissolved crystals. The crystals were frozen as described for the MED7/MED21 complex. After complete exchange of the crystallization to the cryo-solution the crystals were soaked with individually purified MED10. Presence of monomeric protein should ensure complete occupancy of MED10 in the crystal. MED4/MED7 Δ N Δ C/MED21 crystallized with lower concentrations of sodium chloride as well as in conditions containing ammonium acetate. The complex of MED4/MED7 Δ N Δ C/MED9/MED21 did not crystallize. In order to analyze the crystals, they were washed 4 times with the mother solution and dissolved. SDS PAGE analysis of crystals containing MED4 in addition to MED7 Δ N Δ C/MED21, confirmed the presence of all three prmonerico-teins. Crystals were frozen in presence of monomeric MED4, in the same way as the MED7/MED10/MED21 complex described above.

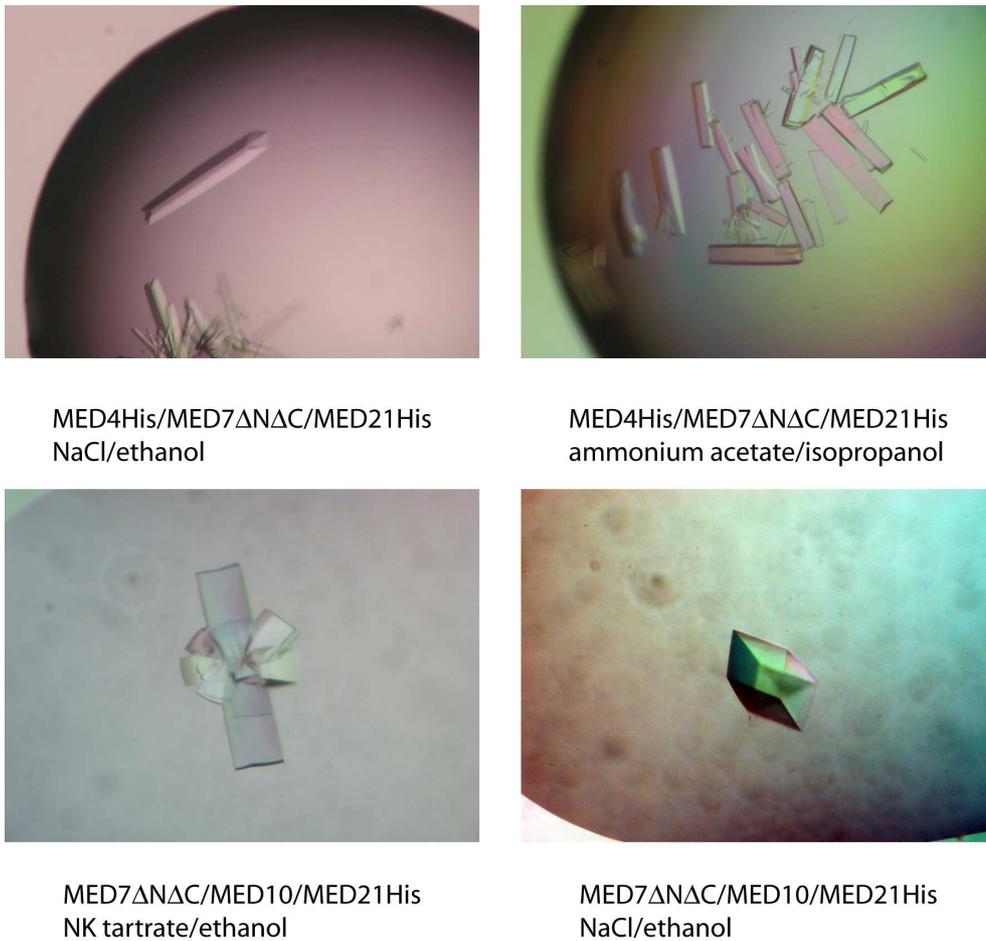


Figure 14: Refined crystals of MED4/MED7 Δ N Δ C/MED21 and MED7 Δ N Δ C/MED10/MED21. The crystallization conditions are listed in Experimental Procedures 12. These crystals were used for structure determination.

Structure determination MED7 Δ N Δ C/MED10/MED21 and MED4/MED7 Δ N Δ C/MED21

Despite the different crystallization conditions diffraction analysis for both trimeric complexes revealed unit cells having the same dimensions as the MED7 Δ N Δ C/MED21 crystals. I solved the structures by molecular replacement in the case of MED7 Δ N Δ C/MED10/MED21, using the refined MED7 Δ N Δ C/MED21 structure as a search model. For MED4/MED7 Δ N Δ C/MED21 The structure was solved by SAD phasing using a selenomethionine-substituted crystal diffracting to 3.0 Å resolution. The structure of MED4/MED7 Δ N Δ C/MED21 was refined to a free R-factor of 31.4 % with data extending to 3.0 Å resolution.

Electron densities of MED7 Δ N Δ C/MED10/MED21 and MED4/MED7 Δ N Δ C/MED21 crystals did not reveal the additional subunits. Despite careful inspection of the electron density, there is no additional electron density corresponding to the presence of the third subunit. This came as a surprise since both trimeric complexes seem very stable during purification, and since analysis of the crystal contents revealed the presence of the additional subunits. Structure determination by selenomethionine labeling and *de novo* SAD phasing for the MED4/MED7 Δ N Δ C/MED21 complex prevents model bias, repositioning of the model in the electron density and refinement confirm the results obtained during the structure determination of MED7/MED21 complex, but do not show any additional electron density. The reason for this could be, that there are only peptide interactions between the MED7/MED21 complex and MED4 or MED10. These interactions might be mediated through less stable regions of the MED7/MED21 complex like the loop between helix α 1 and helix α 2 of MED21. MED4 or MED10 might still be there in the crystal but not visible because MED4 and MED10 are not involved in crystal contacts and therefore not bound in a stable conformation. In agreement with this scenario, the presence of MED10 in the MED7/MED21 complex stabilizes the MED7/MED21 complex during protease digests. In the same experiment MED10 is not stable, indicating that large parts of the protein are not folded. To obtain the structure of larger complexes of the middle module, additional subunits should be included, which could stabilize MED4 and MED10 and reduce their flexibility.

2.2 Reconstitution of a hexameric Mediator middle module that binds RNA polymerase II

Our results show that it is possible to reconstitute subcomplexes of the yeast Mediator middle module by coexpression in *E. coli*. Coexpression can lead to soluble protein complexes. The expressed proteins fold and assemble into their native complexes. To achieve the goal of the reconstitution of a homogeneous Mediator middle module I chose a strategy that combines coexpression of soluble subcomplexes and assembly of these complexes to form a complete middle module. Based on my coexpression studies and the urea washes done by Han et al. (Han et al., 2001), I chose to coexpress MED7, MED10, MED21 and MED31 together. MED7, MED10 and MED21 are expressed from one tricistronic plasmid and a different plasmid harboring MED31His is cotransformed. After purification the heterotetrameric complex can be assembled with separately purified MED4/MED9 to form the middle module.

2.2.1 Purification of a tetrameric MED7/MED10/MED21/MED31 Mediator middle module complex:

Coexpression of MED7/MED10/MED21/MED31 works very well. To selectively purify a stoichiometric complex, versions with a Histidine affinity tag on MED10, MED21 or MED31 were tested. Affinity purification by MED10 or MED21 lead to aggregates of incomplete subcomplexes. However, the tetrameric complex can be purified by a his tag on MED31. A surplus of MED31 can be removed on the Ni column, if bound protein is eluted with a stepwise imidazole gradient. The elution at 70 mM imidazole contained MED7/MED10/ MED21/MED31His. Higher concentrations of imidazole removed residual MED31 from the column, but no stoichiometric complex. Further purification by anion exchange chromatography resulted in two peaks (Figure 15A). Only one contained pure MED7/MED10/MED21/ MED31 as is shown by SDS-PAGE analysis. Separation of the sample by gel filtration results in one single peak, indicating pure and homogenous MED7/MED10/MED21/MED31 complex (Figure 15B).

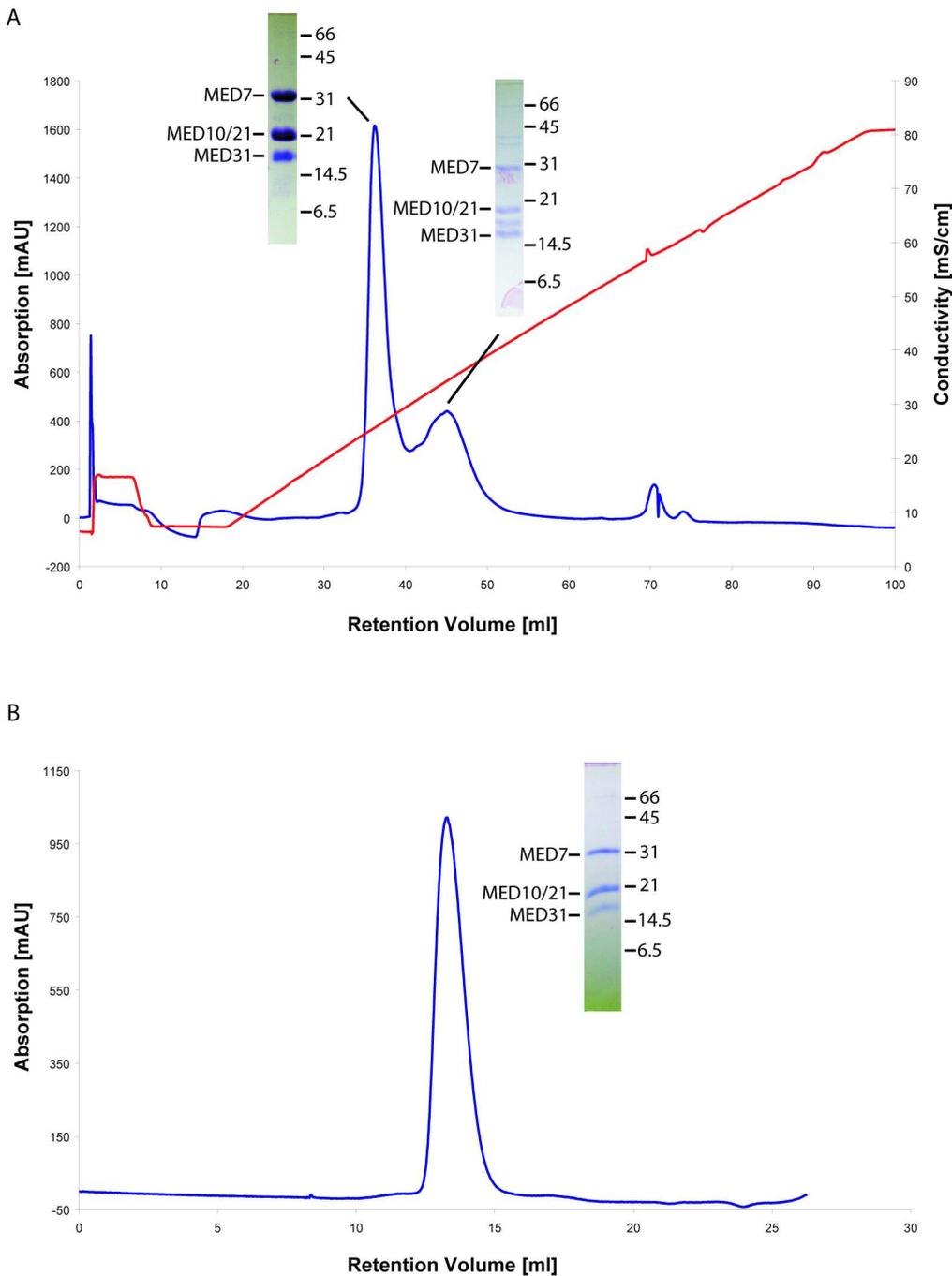


Figure 15: Purification of MED7/MED10/MED21/MED31: (A) Chromatogram of the MonoQ anion exchange chromatography of affinity purified MED7/MED10/MED21/MED31. (B) Superose 6 gel filtration chromatography of the main peak from (A). The absorption at 280 nm is measured to detect protein elution (blue, (A and B)) as well as the conductivity of the elution buffer to follow the salt gradient (red) in (A). SDS-PAGE analysis of the peak fractions is shown next to the corresponding peaks. The names of the proteins as well as the molecular weight (kDa) are given.

2.2.2 Assembly of the hexameric Mediator middle module consisting of MED4/MED7/MED9/MED10/MED21/MED31

For the reconstitution of the six-subunit middle module, MED4 and MED9 were copurified as a stoichiometric complex by ammonium sulfate precipitation (Experimental procedures 4.2.3). The two complexes of MED4/MED9 and Ni affinity purified MED7/MED10/MED21/MED31His (see 2.2.1) were then mixed with a 1.3 fold excess of MED4/MED9. The mixture was assembled by incubation for 1 h at room temperature. The assembled middle module was further purified by anion exchange chromatography. Hexameric middle module elutes at higher salt concentrations from the MonoQ than the tetrameric complex. Therefore free tetrameric complexes are removed (Figure 16A). The excess of MED4/MED9 is removed by a subsequent Superose 6 gelfiltration column (Figure 16B, C). The main peak contains the six proteins in stoichiometric amounts, as estimated from the intensity of the Coomassie Blue-stained bands (Figure 16C). Repeated separation of the main peak on a Superose 6 gelfiltration column reveals one single peak (Figure 16B, D). The retention volume corresponds to a 500 kDa protein complex, which is much bigger than the theoretical molecular weight of 138 kDa. This is probably due to an unusual elongated shape of the complex as was observed for the MED7/MED21 structure. A very elongated structure is also expected from EM data (Asturias et al., 1999). The presence of all proteins was verified by mass spectrometry and Edman sequencing.

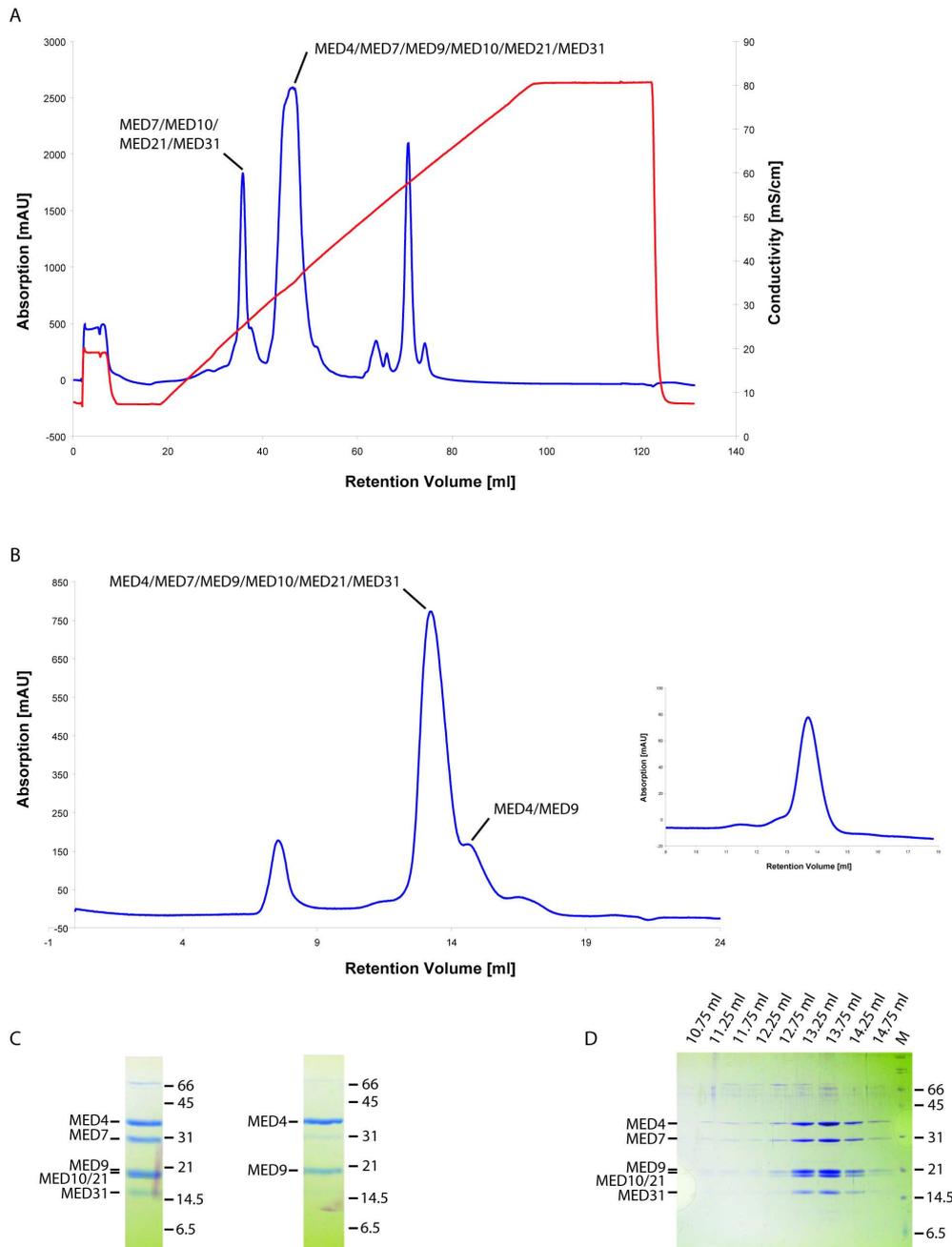


Figure 16: Purification of MED4/MED7/MED9/MED10/MED21/MED31: (A) Chromatogram of the MonoQ anion exchange chromatography of the assembled complex. Complexes containing MED4/MED9 elute later during the salt gradient than tetrameric complexes. (B) Separation of surplus MED4/MED9 from the hexameric complex (left) by gel filtration chromatography. Repeated gel filtration of the hexameric complex results in one single peak (right). The absorption at 280 nm is measured to detect protein elution (blue, (A and B)) as well as the conductivity of the elution buffer to follow the salt gradient (red) in (A). (C) Coomassie stained SDS Page analysis of the Superose 6 runs shown in (B) Proteins and molecular weight standards are indicated.

The resulting complex was subjected to partial proteolysis with trypsin and chymotrypsin. The slow degradation revealed a complex which is stable for at least 10 minutes. This is longer than observed for most subcomplexes. Appearing fragments were identified by Edman sequencing and an overview of identified domains can be seen in Figure 17C and Table 5.

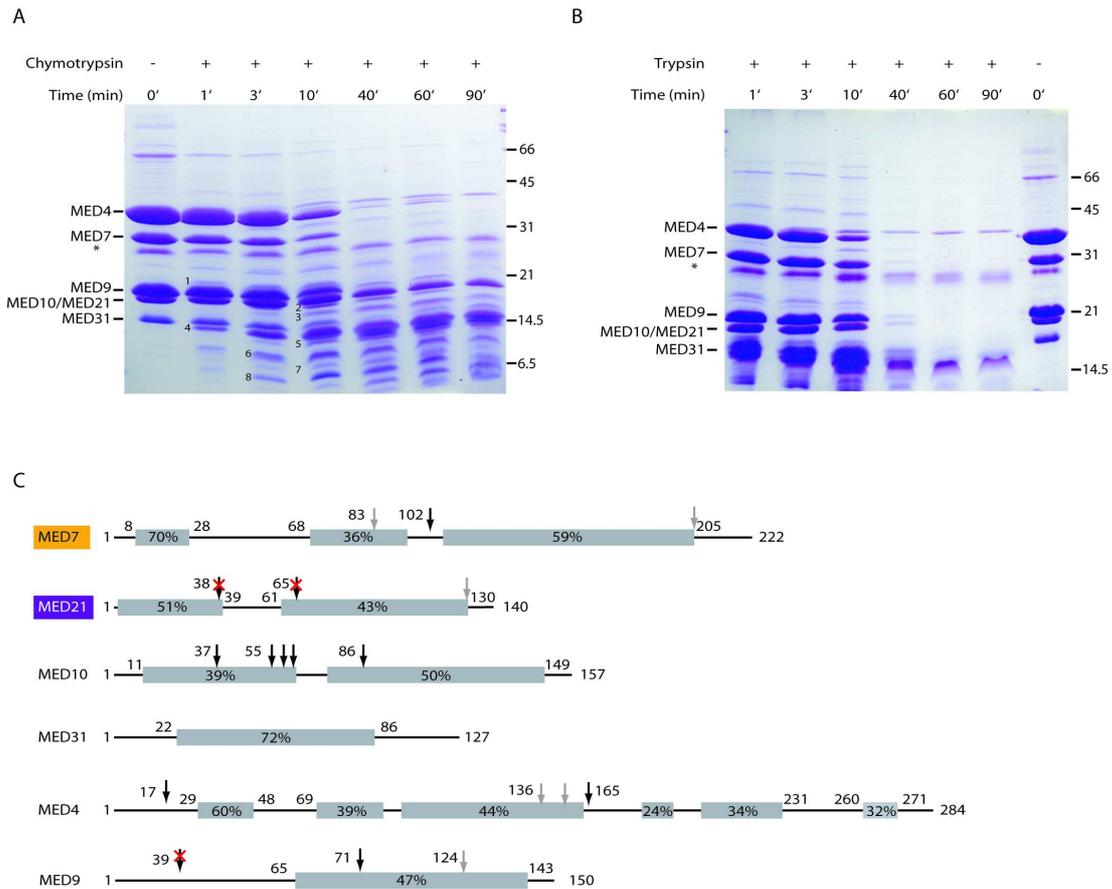


Figure 17: Partial proteolysis of the Mediator middle module: (A) Coomassie-stained SDS PAGE of a chymotrypsin digest of the hexameric middle module complex. Numbers correspond to bands cut and sequenced by Edman sequencing. Corresponding sequences are found in table 5. (B) Coomassie stained SDS-PAGE of a trypsin digest of the hexameric middle module complex. Proteins and Molecular weight are indicated (kDa). The time course of the digests are depicted above the gels. (C) Schematic diagram of the middle module proteins. Grey bars indicate regions of homologous sequences in different species. Numbers correspond to the percentage of homology between yeast and human according to the alignments of MED4, MED10 and MED31 published by Boube et al. (Boube et al., 2002). Proteolytic cleavage sites are indicated above the schemes in black (sequenced) and grey (estimated from SDS-PAGE). Red crosses indicate protection of the site by interaction partners within the middle module compared to the stability of the MED21 alone or MED9 in complex with MED4 only.

Fragment	Sequence	Protein	Assignment (AA residues)	Molecular weight (kDa)
1	MNLQN	MED9	1 to 150	17.4
	TKIPP	MED4	165 to 284	13.4
2	TDRLT	MED21	2 to 140	16.1
	SMGHI	MED4	17 to 164	17.6
3	not identified			
4	KKSTENE	MED7	102 to 222	14.3
	ASNDPGN	MED7	2 to 101	11.3
5	ASNDP	MED7	2 to 83	9.2
6	KLNVR	MED10	62 to 157	11.1
	MNLQN	MED9	1 to 71	8.0
7	IEDGR	MED10	87 to 157	8.3
	ASMNG	MED10	1 to 61	6.8
8	SLHQI	MED9	72 to 124	5.9

Table 5: Partial proteolysis of the Mediator middle module: identified fragments of the chymotrypsin digest. The numbers correspond to the fragments identified in Figure 17.

The MED7 Δ N/MED21 complex which was stable during previous limited proteolysis experiments, is also stable in the middle module complex as expected. The previously unstable N-terminal domain of MED7 however, is protected by additional interactions (Bands 4 and 5; Figure 17) and becomes stable in the complex. Therefore MED7 consists of two domains connected by a protease-sensitive loop. MED10 behaves similarly, having a highly protease sensitive N-terminus and a C-terminal domain. The protection of the C-terminal domain is only dependent on MED7 Δ N/MED21, as limited proteolysis of MED7 Δ N/MED10/MED21 gives similar results. MED31 consists of only one highly conserved domain (70 % between yeast and human, Figure 17C). This domain is stable in all proteolysis experiments done, except for the his tag. The extreme N-terminus of MED31 is required for the stability since N-terminal truncations lead to largely insoluble protein. MED31 is the only protein of the middle module which does not need interaction partners for its folding and stability. Proteolysis only reveals two fragments of MED4. Fragment 2 (Table 5; Figure 17) starts close to the N-terminus of the protein and is similar to fragments obtained in other experiments. Fragment 1 is a C-terminal 13.4 kDa fragment according to N-terminal sequencing. It runs at the position of a 21 kDa protein standard. This strange behavior possibly stems from the many negative charges the domain carries (25 % negatively charged D, E; the-

oretical pI 4.26). A similar behavior is seen for the VP16 activation domain: a strongly charged 7.5 kDa polypeptide (28 % negatively charged D, E, theoretical pI 3.46) which runs at the size of the 21 kDa protein standard (Donaldson and Capone, 1992). These proteolysis results reveal two domains of MED4, both are stable. The C-terminus of MED4 strengthens its interaction with MED21 since truncations of MED4 (residues 1 to 208, 1 to 244) still interact with MED21 but the two proteins are separated by anion exchange chromatography.

Only the C-terminal region of MED9 is conserved from yeast to human (Appendix). However, only the N-terminal part stays intact during proteolysis. Nevertheless the extreme C-terminus of MED9 is needed for stability of the protein since a C-terminal truncation (residues 1 to 134) leads to largely insoluble protein (not shown). The instability of MED4 and MED9 compared to the rest of the middle module might be due to the missing MED1 subunit. The three subunits (MED1, MED4 and MED9) form the link to the tail module of the yeast Mediator (Kang et al., 2001; Guglielmi et al., 2004).

The proteolysis revealed a middle module complex that is stable for at least 10 min when treated with chymotrypsin and trypsin. There are no large unstructured regions, indicating a correctly folded middle module complex. None of the individual proteins MED4, MED10, MED7/MED21, MED4/MED9, except MED31 are stable for this time. Taken together a stable six-subunit middle module complex can be obtained in amounts, purity and homogeneity that are suitable for structure determination experiments.

2.2.3 The Mediator middle module binds Pol II

The Pol II/Mediator middle module complex formed by the 12-subunit Pol II and the six-subunit Mediator middle module was reconstituted using the following approach: endogenous 10 subunit Pol II, the additional recombinant polymerase heterodimer Rbp4/Rbp7 and the recombinant Mediator middle module were mixed in a 1:5:3 molar ratio. The mixture was incubated at 20° C for 1 h and then applied to a Superose 6 gelfiltration column. The column separates the 18 subunit Pol II/Mediator middle module complex from excess Rbp4/Rbp7 and Mediator proteins (Figure 18A). The peak fractions which contain the 18-subunit Pol II/Mediator middle module complex, have a retention volume between the standard proteins thyroglobulin (669 kDa) and ferritin (440 kDa) and correspond to 630 kDa apparent molecular weight. Excess Rbp4/Rbp7 does not coelute with excess hexameric middle module. Thus binding of Pol II and Mediator proteins does not seem to be mediated through direct binding to Rbp4/Rbp7. Fractions containing the holoenzyme were pooled. When the pooled fractions were loaded for a second time on the same gelfiltration column the Pol II/Mediator complex again eluted at the same retention times in stoichiometric amounts, showing that the complex is biochemically stable (Figure 18B).

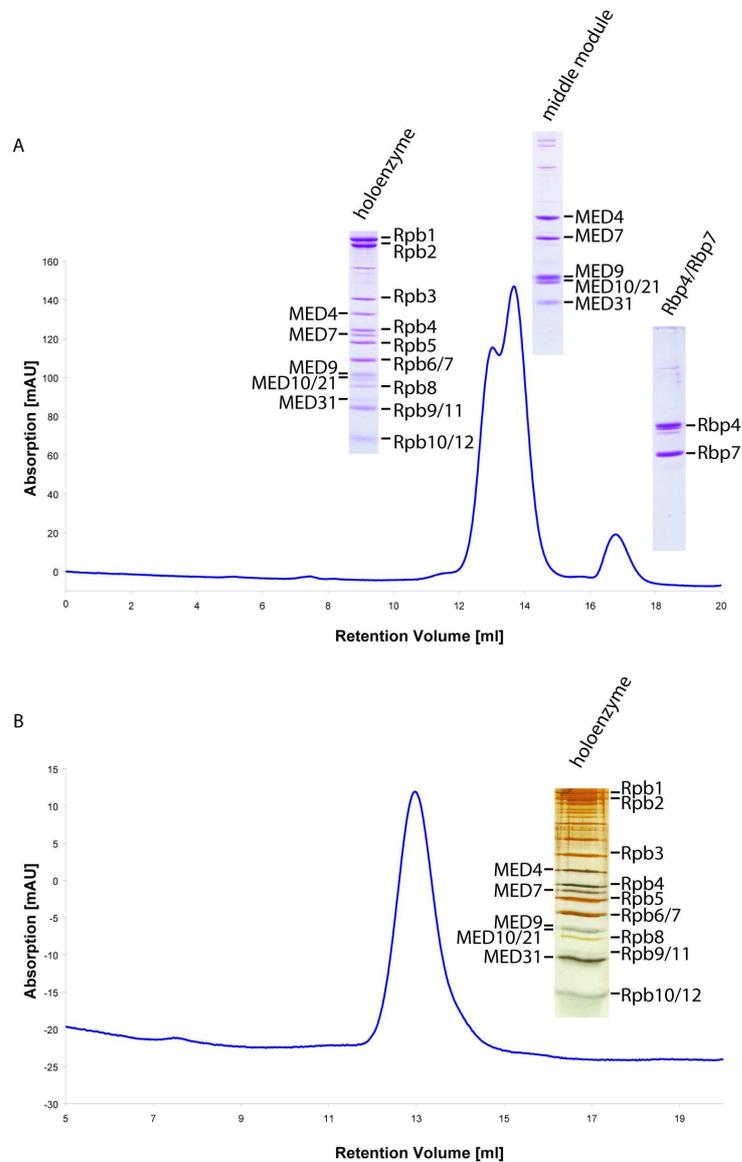


Figure 18: Hexameric middle module binds Pol II: (A) Superose 6 gel filtration of the assembled middle module/Pol II holoenzyme. Coomassie stained SDS-PAGE of the peak fractions are shown next to the corresponding peaks. (B) Second Gelfiltration of the peak fractions from (A). Silver stained SDS-PAGE shows the contents of the peak. Elution of the protein is monitored by the absorption at 280 nm.

Several groups have shown that human as well as yeast Mediator binds to the C-terminal domain (CTD) of the largest subunit of Pol II *in vitro* (Naar et al., 2002). The Srb proteins (Srb2, 4, 5, 7, 8, 9, 10, 11 corresponding to MED20, MED17, MED18, MED12, MED13, CDK8 and cyclinC) were discovered as suppressors of a cold sen-

sitive phenotype of CTD truncation mutants, pointing to possible Mediator/CTD interactions *in vivo* (Nonet and Young, 1989; Thompson et al., 1993). I wanted to test whether the interaction of Pol II and the reconstituted middle module is due to the CTD of Pol II. I therefore performed Gst pull-down assays using a recombinant Gst-CTD (Rpb1 residues 1556-1733) and A Gst-linker-CTD (Rpb1 residues 1437-1733) (Figure 19A). The linker corresponds to a region of the large subunit of Pol II which connects the structured core to the flexible CTD. No interaction was observed of middle module and Gst-CTD or Gst-linker-CTD, while the CTD interaction domain of PCF11 (Meinhart and Cramer, 2004) did interact with both CTD variants, providing a positive control. (Figure 19). This indicates an interaction of middle module with the core Pol II rather than with its CTD. Electron microscopy of the holoenzyme also indicates interactions of Mediator and Pol II in addition to interactions with the CTD (Asturias et al., 1999; Davis et al., 2002).

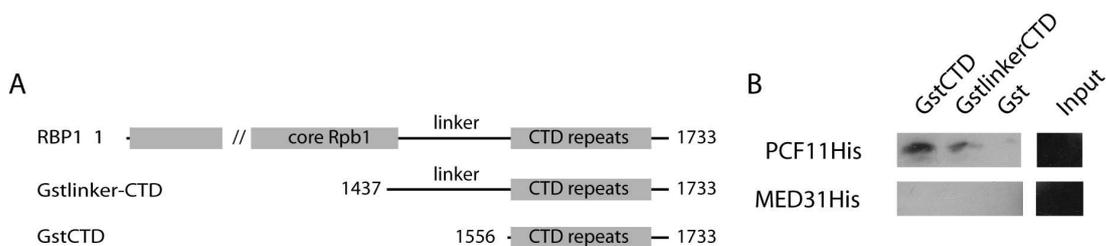


Figure 19: CTD interaction: (A) Scheme of the CTD constructs used. (B) Western blot of the Gst CTD pull down with PCF11 and the hexameric middle module. Detection was done by anti His antibody.

2.3 Structural studies of MED25, a target for the acidic activator VP16

The MED25 subunit of Mediator is unique to higher organisms (Mittler et al., 2003; Yang et al., 2004). It contains two domains. The *von Willebrand domain A* (VWA) binds to human Mediator, and an activator interaction domain (ACID) which binds the activation domain (AD) of the viral protein VP16. There is only one other protein known to share the ACID domain. PTOV1 (prostate tumor overexpressed gene) consists of two copies of the ACID. Figure 20 shows the domain organization, sequence alignment and secondary structure prediction (Rost, 1996) of the ACID. The design of variants and mutations used for crystallization is based on this alignment.

2.3.1 Purification of ACID

Initial ACID variants comprising the human (hACID, residues 393-548) or the drosophila (dACID, residues 515-687) domain are massively overexpressed in *E. coli* and were purified by a C-terminal His tag. hACID has a pI of 9.8 and copurifies with *E. coli* nucleic acids from the expression strain. Purification by Ni NTA and high salt washes of 1 M NaCl removed bound DNA. Subsequent purification by cation exchange chromatography is selective for ACID which is not bound to DNA and resulted in one single peak corresponding to the human ACID domain (Figure 21A). Despite the high pI of 8.9, the drosophila ACID domain did not bind to cation exchange columns. Instead it was purified by anion exchange chromatography, which removes the contaminants as well as DNA. In this case the flow-through of the column contained pure dACID. Gelfiltration of the ACID containing fractions resulted in single peaks, indicating a single and homogenous species of the ACID domains. SDS-PAGE analysis of the peak fractions confirmed the purity of the Protein (Figure 21). Pure and homogeneous material is a prerequisite for crystallization. To ensure complete removal of nucleic acid contaminations in the protein preparation the absorption at 260 nm was monitored during all purification steps. A characteristic of the ACID domain is the yellow color of the purified protein solution (Figure 21B). I monitored the yellow color during the purification by measuring the absorption at 400 nm. The yellow peak coelutes with the protein (280 nm) during all steps of the purification. This color is not due to bound contaminants as shown by SDS-PAGE. ICP (Induced coupled plasma) analysis of the sample did not reveal any metal binding. The pure proteins were subjected to limited proteolysis and proved to be very stable except for a short C-terminal truncation of

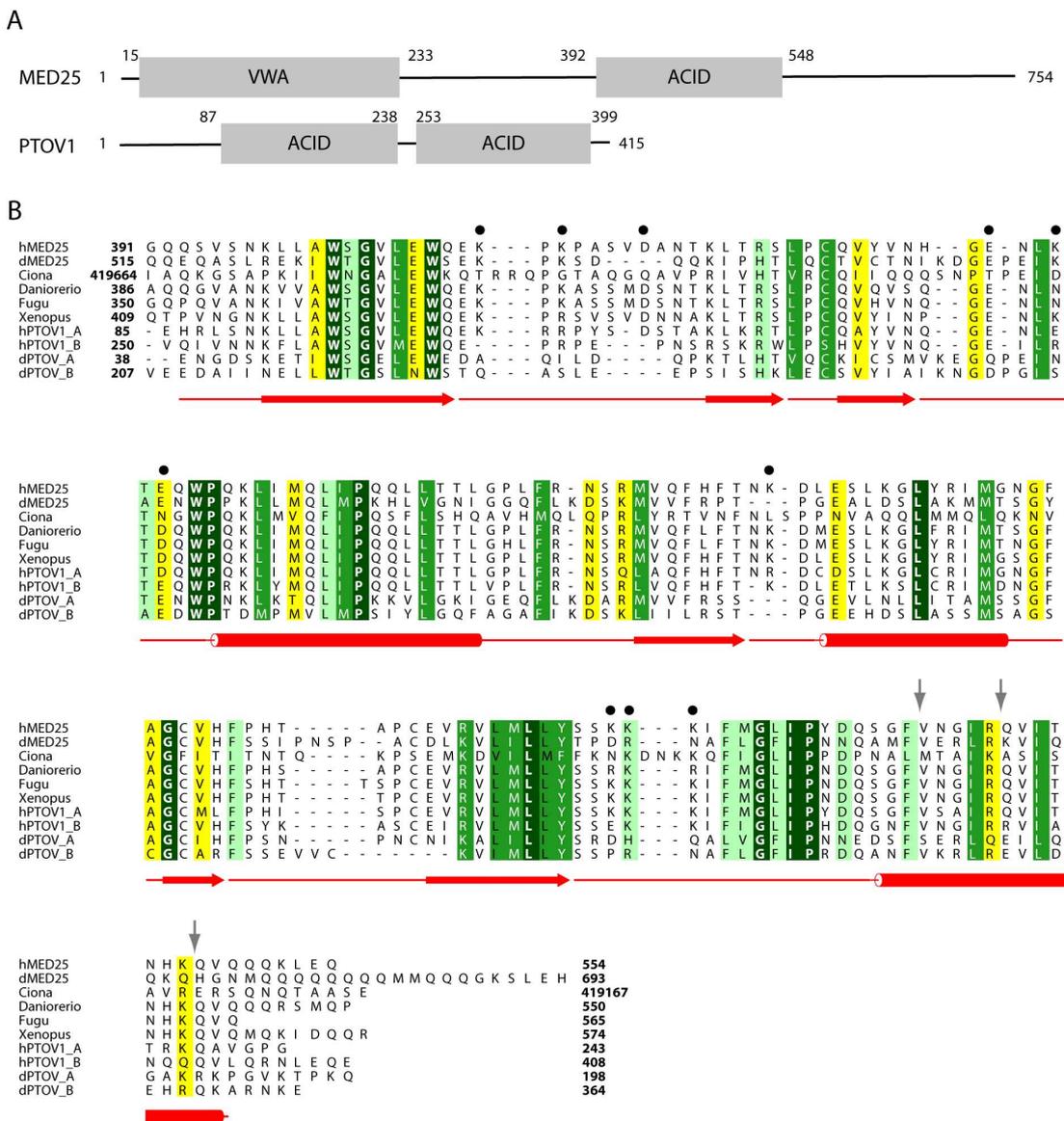


Figure 20: Primary structure of MED25: (A) Schematic representation of human MED25 and PTOV1. The Mediator binding von Willebrand domain A (VWA) and the activator binding domain (ACID) are indicated. PTOV1 consists of two copies of the ACID domain. (B) Multiple sequence alignment of the ACID of MED25 of different species as well as of the two ACIDs of PTOV1 (human and drosophila). Species are indicated in front of the alignment. Predicted strands and helices are shown below the alignment as arrows and cylinders, respectively. Residues are highlighted in dark green, green, light green and yellow according to decreasing degree of conservation. Dots above the alignment indicate residues chosen for mutation to alanine, arrows correspond to estimated protease cleavage sites in the dACID.

dACID (summarized in Figure 20B).

Coexpression of the ACID domain with the complete activation domain (AD) of VP16 (residues 404-490) or VP16H1 (residues 404 to 451) in *E. coli* leads to an increase of protein expression, possibly because the unstructured AD binds to ACID and is thus protected from proteolysis. The proteins could be purified as a stoichiometric complex by a N-terminal Gst tag on the VP16AD (Figure 21C). A second purification step by the C-terminal His tag of the ACID domain results in pure, stable and stoichiometric complexes. Purification by cation exchange or heparin chromatography disrupted the complex and was therefore omitted.

2.3.2 ACID interacts with DNA

Most Mediator proteins of the head and middle domain have acidic pIs. The very high pI of ACID was therefore surprising. To test whether the binding to nucleic acids seen during purification was due to the strong charges of missfolded aggregates of ACID, we tested the DNA binding ability of the purified protein. Band shift assays with double stranded DNA of 19, 21, 23 and 25 bp show a shift of the DNA bound to ACID (Figure 21D). No shift is observed for a shorter double stranded DNA fragment of 17 bp. However, it is not clear whether DNA binding does point to a special role of MED25 in the Mediator complex or it is only due to the strong positive charges on the protein surface.

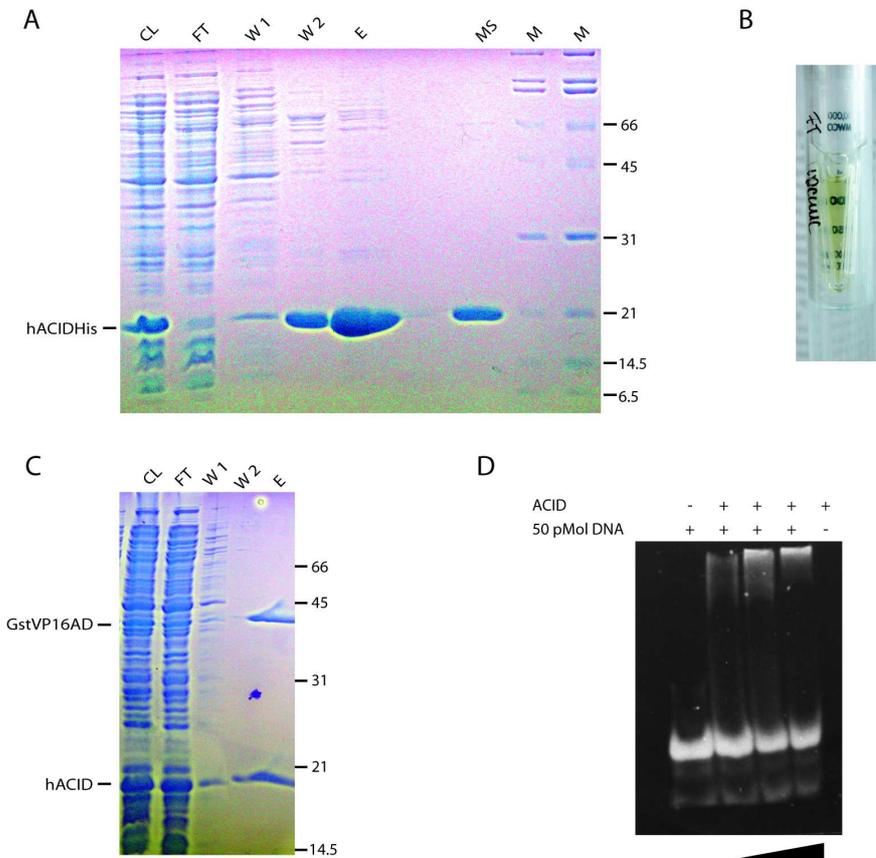


Figure 21: ACID purification: (A) SDS-PAGE analysis of the ACID purification by Ni-NTA (CL: cleared lysate, FT: flow-through, W1, 2: Washing steps, E: elution) and MonoS (MS) cation exchange chromatography. Sizes of the molecular weight standards are given next to the gel, as well as the position of ACID. (B) ACID at a concentration of ~5-10 mg/ml has a light yellow color. (C) SDS-PAGE analysis of ACID/VP16AD purification by glutathione sepharose. Samples are indicated as in (A). The molecular weight is indicated on the right. (D) Bandshift assay of purified ACID with double stranded DNA and increasing amounts of protein (50 pMol, 100 pMol, 200 pMol).

2.3.3 Mapping of a minimal VP16 activation domain

The C-terminal activation domain of VP16 contains two regions, which are responsible for activation. These have been determined by mutational studies (Walker et al., 1993). They interact with multiple components of the transcription machinery (Hall and Struhl, 2002). The activation domain is unfolded in solution (Donaldson and Capone, 1992) and becomes structured upon interaction with the target proteins (Ue-

sugi et al., 1997; Shen et al., 1996; Jonker et al., 2005). For crystallization it is therefore important to determine the exact region of the AD, that interacts with ACID.

The complex of ACID/VP16AD was subjected to partial proteolysis to probe for unstructured regions. Treatment with proteinase K lead to two defined fragments within the activation domain as determined by SDS-PAGE (Figure 22). The N-termini of the two fragments were sequenced by Edman degradation. This revealed a N-terminal fragment of the AD as well as a fragment starting at amino acid 438. The faster migration of the fragments compared to the full activation domain indicates a C-terminal truncation. Because VP16AD consists of two parts with the potential to activate transcription and the end of the first part is after amino acid 451 of the AD, this C-terminal truncation site was estimated to be at position 451. A variant consisting only of the N-terminal H1 region of the activation domain still interacted with the ACID domain, confirming the estimated end of the interaction site. The resulting minimal peptide corresponds to amino acids 438 to 451 of the VP16 AD (Figure 22). This is in agreement with the minimal region for activation as it was determined by (Seipel et al., 1994).

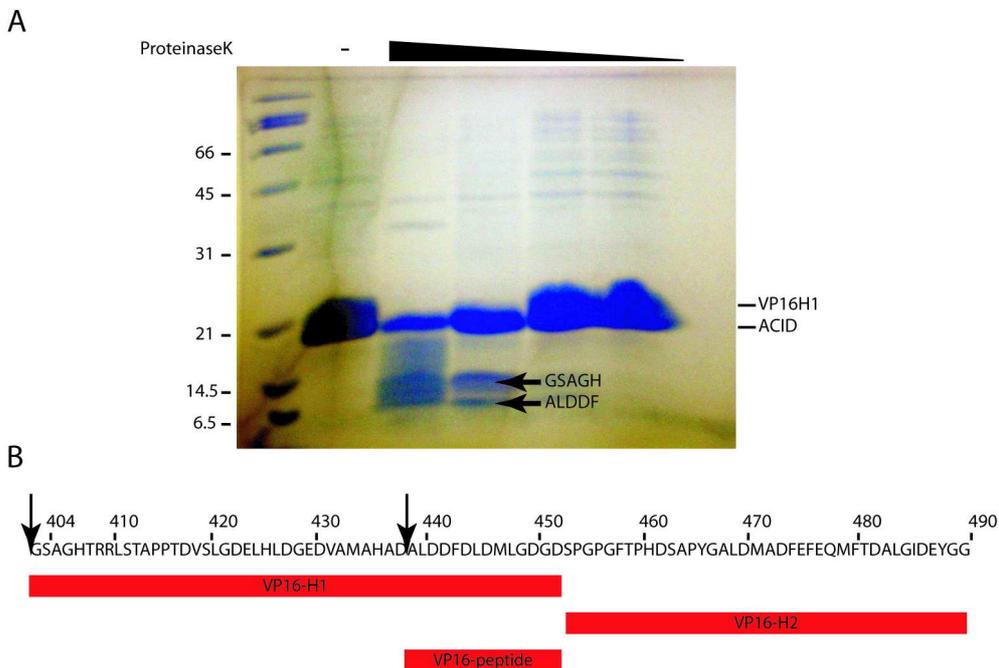


Figure 22: Proteolysis of ACID/VP16AD complex: (A) SDS-PAGE of a proteinase K protease digest of a copurified complex of ACID/VP16AD. The names of the proteins as well as the sizes of the Molecular weight standards are indicated. Sequences determined by Edman sequencing are indicated by arrows next to the corresponding bands. (B) Schematic representation of the sequence of the VP16 AD. Proteolytic sites are indicated by arrows above the sequence. Red bars below the sequence show the fragments corresponding to the identified fragments.

2.3.4 C-terminal truncation of *D. melanogaster* ACID prevents interaction with the VP16 AD

During the attempt on structure determination, crystallization setups of the ACID domain of two different species have been made. The C-terminus of the ACID domain of *D. melanogaster* was sensitive to treatment with the proteases chymotrypsin and trypsin. Based on these proteolysis experiments, C-terminal truncations (535-698, 535-681, 535-687, 535-711) of this domain were cloned and tested for their ability to interact with the VP16 AD. All variants except the shortest (535-698) retained their ability to bind VP16H1 but not a mutant thereof (F442P) (Figure 23).

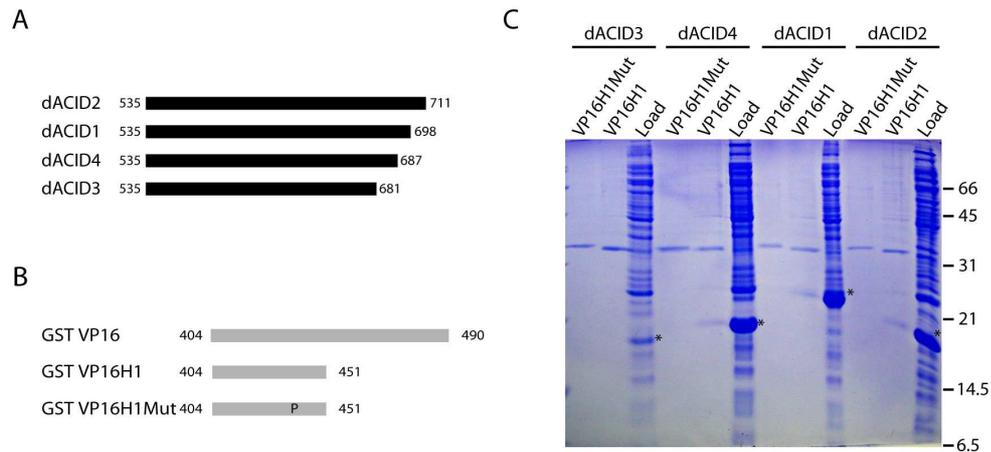


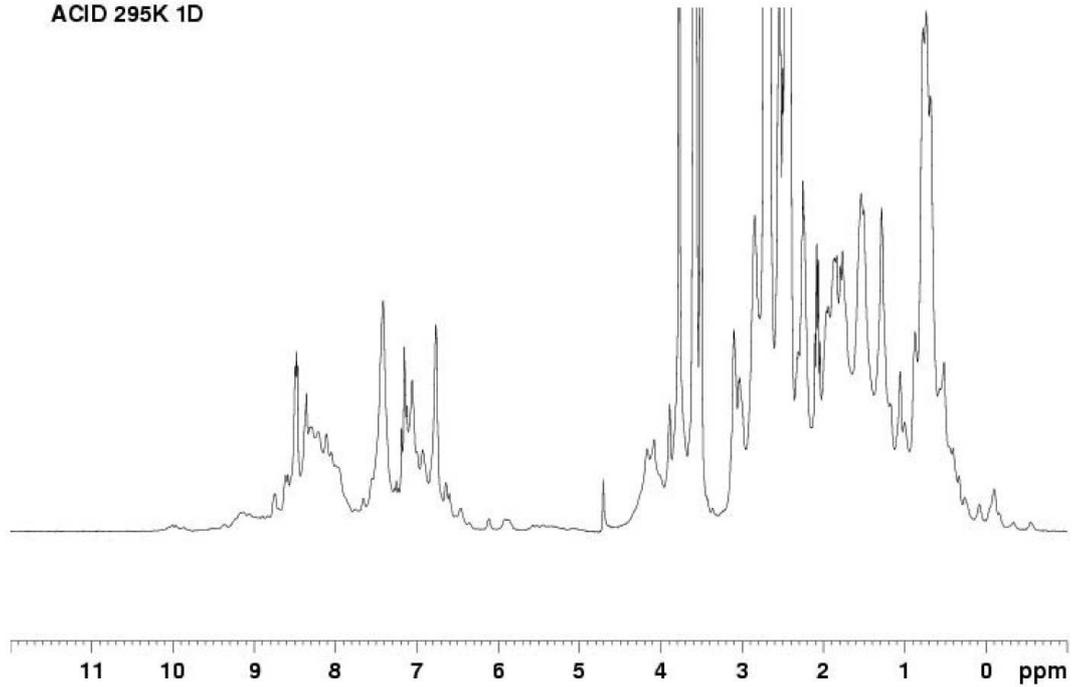
Figure 23: Interaction of dACID with Gst-VP16AD: (A) Schematic overview of the dACID deletion constructs. (B) Schematic overview of the VP16 constructs used the F442P mutant is indicated by P. (C) Coprecipitation of truncated dACID domains with Gst-VP16H1 and with the VP16H1 mutant (F442P). Overexpressed ACID (load, the corresponding bands are marked with *) was precipitated with Gst-VP16H1 and the mutant. Bound proteins were analyzed by SDS-PAGE and Coomassie staining.

2.3.5 Crystallization trials

Several variants of the ACID domain of two different species have been overexpressed in *E. coli* and purified. First, crystallization trials of ACID using 96 well sitting drop plates were set up with the use of a crystallization robot (4.4.6). Since the ACID domain alone did not crystallize, cocrystallization of hACID was tried with a synthetic peptide of the VP16 AD (residues 438 to 451) and with coexpressed VP16H1. However, no crystals were obtained. To ensure folding of the protein, one-dimensional ^1H and two-dimensional ^1H - ^1H and NOESY NMR spectra of the shortest variant of dACID were recorded (Michael Sattler, EMBL Heidelberg). The spectra show that the protein is folded but seems to have also large regions without secondary structure.

The reason that the ACID domain and its VP16 complexes resist crystallization even though it is very well overexpressed, pure and homogenous might lie in the strong positive charges the protein carries. Replacement of putative surface residues having long, flexible and charged side chains by alanine might help the protein to crystallize. The idea is to locally reduce conformational entropy and generate contact forming conformationally homogeneous surfaces. This strategy has been applied by other groups before (Derewenda, 2004). Four variants of the domain were generated

A ACID 295K 1D



B

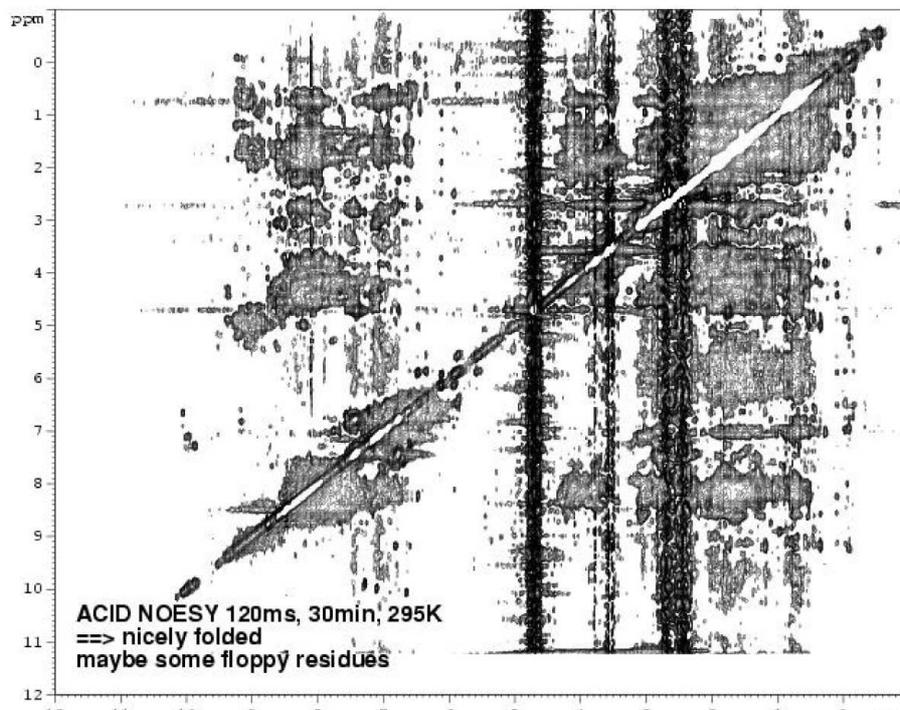


Figure 24: NMR spectra of dACID3 (535-681). (A) One dimensional ^1H NMR-spectra. (B) two dimensional ^1H - ^1H and NOESY NMR spectra. (spectra were recorded by M. Sattler, EMBL Heidelberg)

in which charged residues were replaced by alanines (ACID-K413A/K415A/D420A, ACID-E439A/K442A/E444A, ACID-K480A, ACID-K519/520/521A). All of the mutated residues lie in regions without predicted secondary structure (Figure 20), which are likely to locate on the surface of the ACID domain. All variants were screened for crystallization, unfortunately without success.

3 Discussion

3.1 An *E. coli* coexpression system to analyze pair-wise interactions within the Mediator middle module and to obtain crystallization-grade Mediator subcomplexes

A significant impediment to high-resolution structure determination of yeast Mediator stems from its low abundance within yeast cells and the difficulty to obtain the complex in pure form, free from other components of the transcriptional machinery (Myers et al., 1998; Takagi et al., 2005). Heterologous gene expression of most single Mediator subunits in *E. coli* leads to insoluble or not properly folded proteins, because the correct binding partner is absent. During this study I established a recombinant expression system in *E. coli* which allowed to study protein-protein interactions within the middle module of Mediator. Coexpression of subunits often leads to solubly expressed heterodimeric complexes of interacting subunits.

Using the coexpression approach, I obtained a detailed interaction map between the six well conserved of the seven proteins in the yeast Mediator middle module (Figure 25). This map shows the following interactions: MED4-MED7, MED4-MED9, MED4-MED21, MED7-MED21, MED10-MED21. All of the interacting subunits could be purified after coexpression in a stoichiometric manner over several chromatographic steps, showing the stable nature of the interactions. In addition I showed that MED31 contacts several subunits of the middle module, but needs more than one partner for the formation of a stable complex. It is interesting to note that at least MED7/MED21 can only be obtained in a stoichiometric manner if the proteins are coexpressed. Individual expression and subsequent assembly of the two proteins does not lead to a stable complex. All of these findings are consistent with published data on yeast Mediator subunit interactions, which stem from coexpression of subunits in insect cells (Kang et al., 2001; Koh et al., 1998), coimmunoprecipitation (Lee and Kim, 1998), the split ubiquitin assay (Gromoller and Lehming, 2000a), and from yeast two-hybrid analysis (Uetz et al., 2000; Guglielmi et al., 2004; Ito et al., 2001). These studies additionally showed that the remaining subunit of the middle module, MED1, binds to the MED7 N-terminal region. The Med7 N-terminal region also binds MED9 (Guglielmi et al., 2004; Kang et al., 2001).

MED31 is a highly conserved protein that copurifies with metazoan Mediators. In

yeast, MED31 was genetically associated with transcription for a long time but only recent reports confirmed copurification with *S. cerevisiae* as well as *S. pombe* Mediator complex (Guglielmi et al., 2004; Linder and Gustafsson, 2004; Takagi et al., 2005). Using the coexpression strategy I found interactions of MED31 with MED4, MED7, MED10 and MED21. Therefore MED31 is an integral part of the Mediator middle module. The interactions to MED10 and MED21 were also observed by Guglielmi et al. in yeast two hybrid screens. The interactions to MED4 and MED7 are in keeping with data from a large-scale yeast two hybrid screen (Ito et al., 2001). One possible reason why MED31 was not identified earlier as a Mediator subunit is its size. MED31 may have escaped earlier detection by SDS-PAGE due to its small size of 14.7 kDa. Furthermore small subunits are stained less than their bigger interaction partners and might be missed by visual inspection of the SDS-PAGE.

MED1 is the only protein of the middle module which could not be obtained using the coexpression-copurification method. MED1 is a 64 kDa protein which is only weakly conserved from yeast to human. In yeast multiple interaction partners have been identified by two hybrid analysis (Uetz et al., 2000; Guglielmi et al., 2004; Ito et al., 2001). Beside the binding to MED4, MED9 and MED7 in the middle module, MED1 seems to form a bridge to the tail module: it binds to MED14, as well as MED5 (Guglielmi et al., 2004). Folding of MED1 can be dependent on these proteins or on the cellular environment of eukaryotic cells which is not provided by *E. coli*. Except for MED1, expression of two interacting subunits in one cell could promote folding of the subunits and allowed for their recombinant expression. Subsequent copurification gave an indication about the stability of these interactions. Amounts and quality of the copurified heterodimers were sufficient for structural studies.

3.2 The MED7/MED21 structure

Limited proteolysis identified the non-structured regions within the MED7/MED21 heterodimer and subcloning of the stable part allowed to obtain a crystallizable sub-complex. Initial crystals diffracted to a maximum of 3.3 Å resolution and were very radiation sensitive. MAD data of selenomethionine crystals was obtained to only 3.8 Å and did not allow structure determination. Removal of the protease sensitive C-terminus of MED21 and introduction of additional methionines improved the crystals and enhanced the anomalous signal in selenomethionine derivatives. MAD data

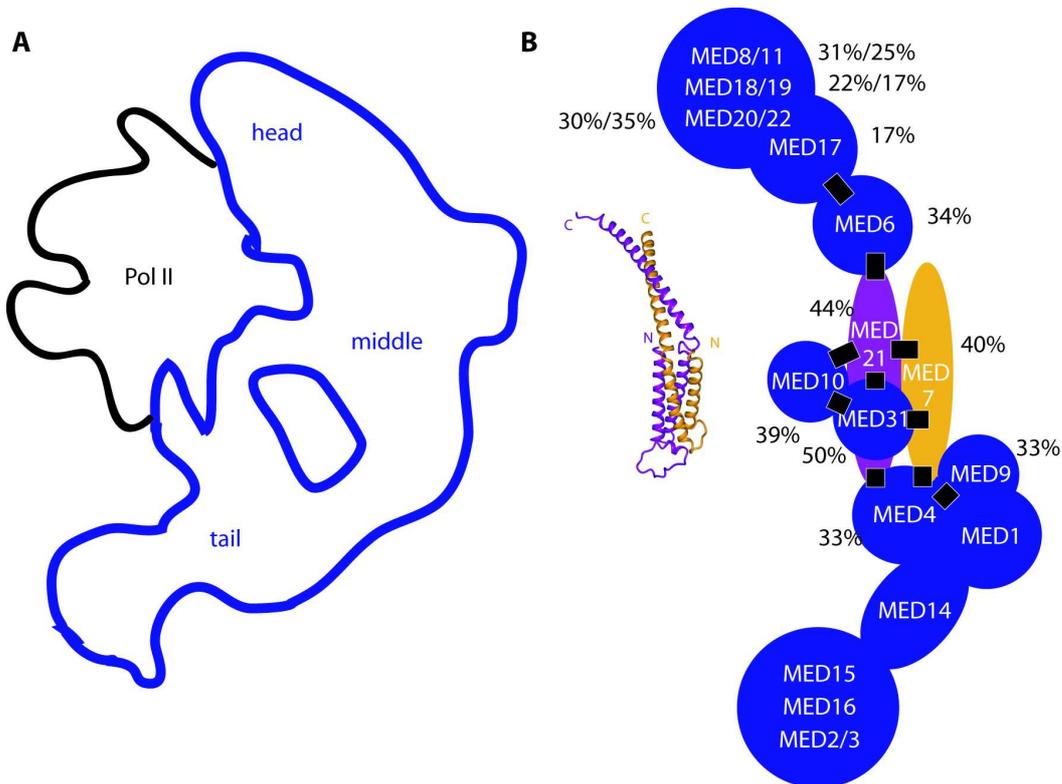


Figure 25: Mediator architecture: (A) Outline of Mediator bound to Pol II according to electron microscopy (Davis et al., 2002). The suggested head, middle and tail modules are indicated. (B) Mediator subunit architecture. All depicted subunits are at least partially conserved throughout eukaryotes except MED2 and MED3. The percentage of sequence homology between yeast and human subunits is indicated for the better conserved subunits. The four-subunit CDK8/CyclinC module and the putative Mediator subunit MED5 have been omitted. For size comparison, the structure of the MED7/MED21 complex and the Pol II-Mediator complex are drawn to scale. Black rectangles indicate direct protein-protein interactions detected in this study.

to 3.5 Å allowed structure solution. The subcomplex contains the conserved parts of MED7 and MED21. The MED7/MED21 heterodimer binds MED6 in the head module and the most conserved subunits (MED4, MED9, MED10, MED31) of the middle module. The interaction mapping established the MED7/MED21 complex together with MED31 as the conserved assembly scaffold of the middle module. The structure of MED7/MED21 shows a very elongated protein complex. Even more elongated structures can be expected for the complexes containing MED7/MED21 and MED4 or MED4/MED9 in addition. Analysis of the recombinant complexes by gelfiltration showed very short retention times. Sequence-based predictions of secondary structure (Rost, 1996) and coiled-coils (Lupas et al., 1991) strongly suggest that subunits MED9, MED10 and MED4 are largely helical and form coiled-coils that may mediate the strong subunit interactions observed here, and may span large distances on the Pol II surface.

3.2.1 Conserved Mediator core architecture

My data show that the MED7/MED21 heterodimer tightly binds to other highly conserved Mediator subunits via its evolutionarily conserved regions and thus plays a central architectural role within Mediator. Analysis of the molecular surface of the MED7/MED21 heterodimer reveals that the majority of the surface is hydrophobic, consistent with the observed extended interactions with other subunits. Four conserved surface patches (Figure 26A) may serve as protein interaction sites and befit the architectural role of the MED7/MED21 heterodimer. Patch 1 and 2 are near the hinge region. Patch 2 includes the highly conserved N-terminal helix $\alpha 1$ of MED21 (residues R4, Q7, L8), and a part of MED21 helix $\alpha 2$ (residues L76, S79, L80, Figure 26A). Since the highly conserved seven N-terminal MED21 residues are required for MED6 binding (Gromoller and Lehming, 2000a), I propose that patch 2 constitutes the MED6 binding site. Patches 1, 3 and 4 may bind to subunits of the middle module, but may also be involved in interactions with Pol II. Patches 3 and 4 are involved in crystal contacts. Patch 3 corresponds to the open end of the coiled-coil that stacks against a neighboring coiled-coil in the crystals (Figure 9B), and patch 4 mediates dimerization of bundle domains of neighboring MED7/MED21 heterodimers in the crystals (Figure 9C). The resulting two types of tetramers in the crystal do not show additional conserved surface patches (Figure 26B, C). Taken together, the high conservation of the MED7/MED21 heterodimer, its many interactions with conserved subunits, and its

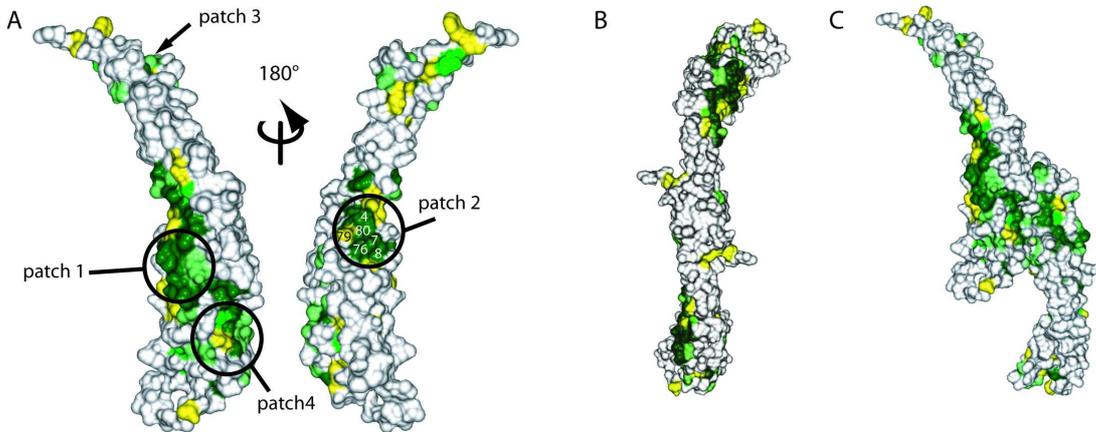


Figure 26: Conserved surfaces of MED7/MED21 (A) Conserved surface patches on the MED7/MED21 complex. The views are as in Figure 9A. Conserved residues are colored according to Figure 9. Four conserved surface patches are indicated. (B) Surface representation of the tail to tail heterotetramer observed in the crystal packing. The view is as in Figure 9B. Conserved residues are colored according to Figure 9 (C). Surface representation of the head to head heterotetramer observed in the crystal packing. The view is as in Figure 9C. Residues are colored in dark green, light green and yellow according to decreasing degree of conservation.

extended conserved hydrophobic surface patches all indicate that the structural architecture of the central region of the core Mediator is the same in all species.

3.2.2 Conserved hinges and Mediator function

Electron microscopy of free yeast Mediator and Mediator bound to Pol II revealed a dramatic structural rearrangement (Davis et al., 2002). Whereas free Mediator is relatively compact, it becomes extended upon Pol II binding, and is wrapped around the polymerase surface in the Pol II-Mediator complex (Figure 25A). This transition apparently involves a large change in the relative position of the Mediator middle and head modules (Davis et al., 2002). Our data suggest that the MED6 subunit plays a central role in this process, because it bridges between two modules of the core Mediator. Secondary structure prediction for MED6 reveals extended loop regions and helices with low probability (Rost, 1996), indicating a strong intrinsic flexibility for MED6. MED6 may form a conserved flexible connection between the head and middle modules. The bridging role of MED6 is relevant for Mediator function *in vivo*. In agreement with the essential bridging role a point mutation in MED17 suppresses

temperature-sensitive mutations in the N-terminal region of MED6 that weaken the interaction with MED17 (Lee and Kim, 1998). The interaction of MED21 with MED6 is apparently essential *in vivo*, as a deletion mutant of yeast lacking the 15 N-terminal residues of MED21, which does not bind MED6 *in vitro*, is not viable (Gromoller and Lehming, 2000b).

Strong structural changes have also been observed in the mammalian Mediator coactivators CRSP and ARC upon binding of activator proteins (Taatjes et al., 2002, 2004b). In addition to the MED6 hinge, the intrinsic flexibility of the MED7/MED21 heterodimer may account for these conformational changes. The hinge region between the two domains of the heterodimer may allow flexibility within the middle module, because the relative locations of subunits bound to surface patch 3 on the coiled-coil and patches 1, 2, and 4 on the bundle domain can change. Binding of MED6 to the hinge region of the MED7/MED21 heterodimer may restrict or coordinate hinge motions. The repositioning of different parts of the core Mediator, enabled by conserved hinges as suggested here, and triggered by the interaction with various partners, may be crucial for Mediator function.

3.3 Larger complexes

I used the crystallized MED7/MED21 subcomplex as a scaffold for the cocrystallization with additional subunits. Despite the very stable nature of MED4/MED7/MED21 and MED10/MED7/MED21 complexes during purification, MED4 and MED10 were not visible in the electron densities. There are, however, two findings indicating the presence of the two additional proteins in the crystals. First, both complexes crystallized easier and also in different conditions than the original MED7/MED21 heterodimer. Second, analysis of the crystal contents by mass spectrometry and SDS-PAGE revealed the presence of MED10 and MED4 in their respective crystals. In addition, MED4 is a 36 kDa protein, bigger than the crystallized MED7/MED21 heterodimer. Therefore more than 50 % of the protein content in the crystallization drop consisted of MED4. Such a high percentage of MED4 in the drop would be likely to inhibit crystal growth of the MED7/MED21 complex.

The invisibility of MED4 and MED10 in the electron densities might be due to high mobility of the proteins within the crystals. Such high mobility could arise if the proteins bind to mobile regions with high B-factors, like the loop between helix $\alpha 1$

and helix $\alpha 2$, or the C-terminus of MED21. Binding of MED10 to the loop region of MED21 is supported also by limited proteolysis experiments where MED10 stabilizes the MED7/MED21 complex, possibly protecting its largest loop region. An answer to this problem might lie in the additional proteins of the middle module. These proteins could stabilize MED4 and MED10 by binding the two proteins and at the same time bind to the surface patches (Figure 26) in the MED7/MED21 heterodimer. Thereby the flexible loops of the MED7/MED21 heterodimer as well as the additional proteins would become stable. A first step in this direction is the reconstitution of the four- and six-subunit middle modules.

3.3.1 A reconstituted Mediator middle module

Structural analysis of the entire Mediator complex has so far been limited to low resolution EM images due to non-homogeneous samples of purified endogenous Mediator/Pol II holoenzyme. These reports were very informative, showing the conformational changes as well as the overall structural conservation of Mediator complexes from yeast to human. However, to understand the mechanism of Mediator as a coactivator, detailed structural information is needed. Such detailed structural analysis requires well defined and homogeneous samples. Because of its modular organization, its low abundance and the many interactions to activators and the transcription machinery, it is difficult to purify endogenous Mediator as a homogeneous complex. Therefore we reconstituted Mediator subcomplexes from recombinant proteins expressed in *E. coli*. Coexpression of single subunits allowed their recombinant expression in soluble form. Assembly of two subcomplexes of two and four proteins lead to the six-subunit middle module. This middle module contains all the highly conserved Mediator core proteins except for MED1 (Figure 27). It is a well defined, homogeneous and stoichiometric complex of known composition and therefore suited for high resolution structural analysis by cryo EM or crystallography. A first quality control by negative stain EM shows single particles of the expected size.

With this complex in hand, a Pol II/Mediator middle module complex could be reconstituted. To my knowledge the reconstituted six-subunit middle module is the first recombinant Mediator subcomplex shown to bind purified Pol II. We now have a tool to answer questions of how Mediator integrates the signals from transcription factors to the polymerase. For instance, it is still unclear which subunits of Mediator contact

the polymerase. Approaches using yeast two hybrid screens did reveal intersubunit interactions of Mediator but none to the polymerase (Guglielmi et al., 2004, personal communication M. Werner). Initial binding studies done with the purified recombinant dimeric and trimeric Mediator subcomplexes (this study) did not bind to endogenously purified Pol II. It is possible that there are no strong contacts between single subunits of the two complexes but each interaction surface is made up of more than one protein, explaining why Mediator-Pol II interactions are not detectable in yeast two hybrid screens. Published EM data show multiple interaction sites between Mediator and Pol II (Asturias et al., 1999). It is likely that there are multiple weak cooperative contacts adding up to the interaction. The interaction of Pol II and Mediator has to be reversible when the Pol II starts transcription and Mediator remains at the promoter. Multiple weak contacts are possible to break by conformational changes, while this is unlikely for a single strong interaction site.

The binding of the middle module to the polymerase and its stable association also during repeated gel filtration chromatography indicates that the *in vitro*-assembled complex is properly folded and reconstituted. No binding of the reconstituted middle module to the CTD was observed. The reconstituted middle module associates with Pol II. Therefore the CTD alone is not sufficient for middle module binding but the core of Pol II is needed for the interaction. However, Kang et al. coexpressed a six subunit Mediator middle module (consisting of MED1/MED4/MED7/MED9/MED10/MED21) in insect cells which could coprecipitate Gst-CTD. Their complex contained additionally MED1 but not MED31. The difference seen in CTD binding might lie in the different subunit composition or post-translational modification of the complexes. *E. coli* does not allow for phosphorylation of the subunits while this is possible in insect cells. It was reported that MED4 is phosphorylated in holoenzyme complexes (Guidi et al., 2004). Different subunit composition or phosphorylation could also induce a different conformation of the module. It was shown by EM that the open (polymerase bound) conformation of Mediator also occurs at high pH or, partially, in the presence of the CTD (Asturias et al., 1999).

3.4 Conservation within Mediator subunits

Mediator has been conserved during evolution. EM studies revealed similarities in the overall structures of yeast and mammalian Mediator complex (Asturias et al., 1999;

Dotson et al., 2000). Identification of orthologues of 22 of the 25 *S. cerevisiae* subunits (all subunits of *S. cerevisiae* except MED2, MED3 and MED5) provide a rationale for the overall structural similarities observed. The limited overall sequence conservation from yeast to human is possibly sufficient for conservation of the folds of the single subunits and their intersubunit contacts.

Most of the Mediator subunits are only weakly conserved from yeast to human. Especially in the tail module only single domains, if at all, have orthologues in yeast and human. Therefore large parts of the surface of the Mediator complex cannot be conserved. This could allow the Mediator to accommodate the huge increase in regulatory inputs during evolution, emerging from novel gene specific transcription factors.

Some subunits however, are stronger conserved than needed for preservation of the overall fold. Moreover, the degree of conservation is not linked to whether these proteins are essential in yeast or not. Together with CDK8/CyclinC, MED31 is the highest conserved Mediator subunit (50 % homology) (Linder and Gustafsson, 2004). None of these proteins are essential for yeast viability (Giaever et al., 2002; Fan and Klein, 1994). A reason for the higher conservation is preservation of enzymatic activity as it is observed for the CDK8/CyclinC module. MED31 together with MED7, MED10 and MED21 are the highest conserved subunits of the core Mediator (Figure 27). This might be because the middle module is the part of Mediator which interacts with the (highly conserved) Pol II as determined in this study. The flexibility observed for the MED7/MED21 complex and predicted for MED6 might require such a high conservation. Indeed, conserved residues of MED7 and MED21 map to the hinge region (Figure 9).

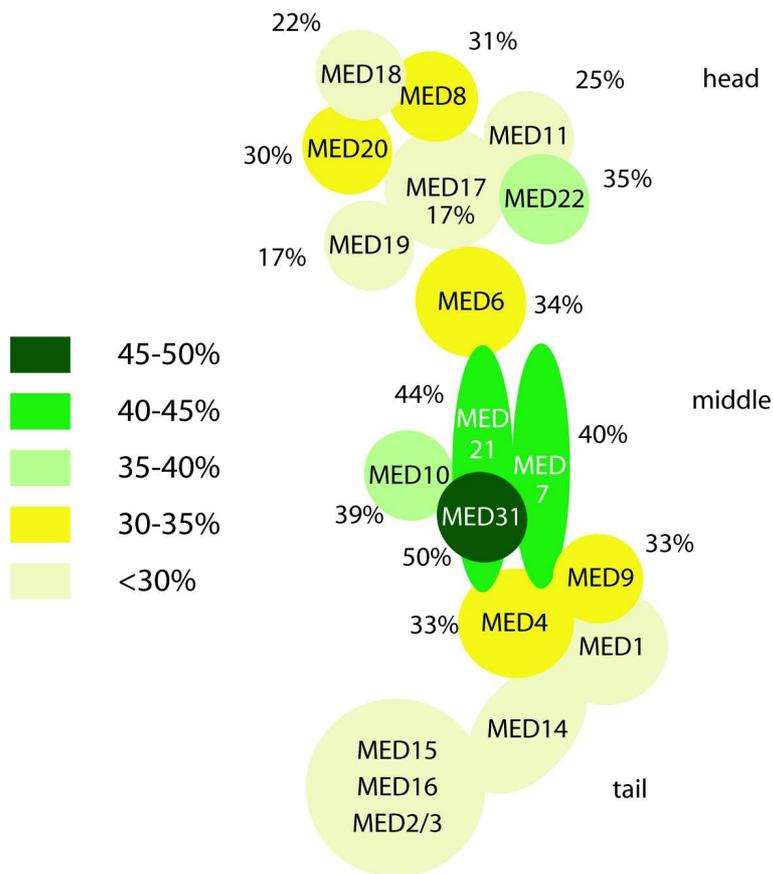


Figure 27: Mediator conservation: the Mediator subunit architecture is shown as in figure 25B. Head middle and tail modules are indicated. All depicted subunits are at least partially conserved throughout eukaryotes, except MED2 and MED3. The percentage of sequence homology between yeast and human subunits is indicated for the better conserved subunits. The color code is given on the left. *S. cerevisiae* has long insertions in the subunits MED6 and MED18, the conservations given are for the whole proteins.

3.5 Towards the structure of an activator interaction domain

The activator interaction domain (ACID) of MED25 is a small 20 kDa protein domain. During this study I established purification protocols for *human* and *drosophila* ACID. Analysis by proteolytic digests revealed the stable nature of the domain. Except for the C-terminus of the *drosophila* protein no degradation was observed, indicating folding of the domain. Recombinant ACID retains its ability to bind to VP16, again pointing to a folded domain. Nevertheless no crystals could be obtained. Therefore I did differ-

ent ACID truncations and mutations. A C-terminally truncated version consisting of dACID amino acid residues 535-681 lost the VP16-binding ability. N-terminal truncation of hACID and recombinant expression did not lead to soluble protein, possibly because the deleted N-terminus is needed for folding. Outgoing from these results I generated the minimal hACID and introduced mutants to remove patches of strongly charged and flexible amino acids in regions without any predicted secondary structure. The mutations to alanine might generate surfaces which potentially could be involved in crystal contacts (Derewenda, 2004).

Purified concentrated ACID has a yellow color. This is not due to metal binding of the protein as shown by ICP analysis. There might be a cofactor, which has a yellow color such as NADH and binds to the protein such that it can not easily be removed by conventional chromatography methods. However, there is no enzymatic activity known, that could explain the presence of such a cofactor.

I found that ACID copurifies with nucleic acids. Electrophoretic mobility shift assays of purified recombinant protein show DNA binding *in vitro*. ACID has a strikingly high pI of 9.7, which differs from most other Mediator subunits. The DNA binding is not only due to the high charges of nucleic acids, since it is also dependent on the length of the DNA used. Together with the notion that MED25 is associated in part with the chromatin fraction of human cells (Blazek, 2005) it might be that the DNA binding observed is not just an artefact which stems from the highly charged domain, but it can indeed associate with nucleic acids. ACID interacts with the activation domain of VP16 (Mittler et al., 2003; Yang et al., 2004). The core of VP16 binds to the DNA and cellular DNA binding factors, like OCT1 and HSF1 (Liu et al., 1999) it is possible that ACID contributes to this interaction.

The protein domain is small, highly soluble and stable and should therefore be an ideal target for crystallization. Despite many efforts to crystallize ACID, I could not obtain protein crystals. Additional stabilization of the domain by the addition of VP16H1 or the mapped VP16 interaction peptide did only lead to two dimensional protein crystals which do not allow for structure determination. Nevertheless NMR tests of the protein solution showed that the protein is folded. Given the small size of less than 20 kDa, and the high solubility it is an ideal target for structure determination by NMR.

3.6 Conclusion and Outlook

During this thesis a recombinant coexpression system in *E. coli* was established, which allowed analysis of the pairwise interactions within the Mediator middle module as well as the generation of crystallization grade subcomplexes. Solution of the structure of the MED7/MED21 heterodimer in two different crystal forms revealed a flexible hinge, which could be responsible for the large conformational changes Mediator undergoes upon polymerase binding. Further, a protocol for the reconstitution of a recombinant middle module has been established. Amounts and quality of the middle module complex are sufficient for structural studies. This recombinant middle module is a minimal Mediator module required for polymerase binding.

Structural analysis of the Mediator middle module will clarify the role of the proposed conformational changes seen in the MED7/MED21 structures. Further research will have to focus on complexes of the Mediator middle module together with the polymerase. Medium resolution EM or crystal structures of the Pol II/middle module complex and subsequent docking of high resolution substructures are necessary to elucidate how Mediator binds and regulates the polymerase. Activators like VP16 bound to single domains of the Mediator will further enlarge our knowledge of the signal transmission from activators to Mediator. A long term goal of the study started here is the *in vitro* reconstitution of a functional core Mediator and its structural analysis alone as well as in complex with Pol II and the general transcription factors.

4 Experimental procedures

4.1 General Methods

4.1.1 Bacterial strains

strain	description	source
DH5 α	F' Δ 80dlacZ Δ M15 Δ (lacZYA-argF)U169 deoR recA1 endA1 hsdR17(r _K ⁻ m _K ⁺) phoA supE44 λ - thi-1 gyrA96 relA1	Woodcock et al., 1989
XL-1 blue strain	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac[F' proAB lacIqZ Δ M15Tn10(Tetr)] constitutive promoter	Stratagene
BL21-Codon Plus (DE3)-RIL	B F- ompT hsdS(rB- mB-) dcm+ Tetr gal λ (DE3) endA Hte [argU ileY leuW Camr], extra copies of argU, ileY and leuW tRNA genes on a ColE1-compatible plasmid with chloramphenicol resistance marker, protease deficiency, chromosomal T7-polymerase gen	Stratagene
BL21-(DE3)	B F- ompT hsdS(rB- mB-) dcm+ Tetr gal λ (DE3) endA Hte, protease deficiency, chromosomal T7-polymerase gen	Stratagene
B834	<i>E. coli</i> (DE3) (hsd metB)	Budisa et al., 1995

Table 6: Bacterial strains

4.1.2 Plasmids

#	Insert	Vector	Restriction sites	Tags	Remarks	Expression level
1	MED10	pET21b	Nde I, Not I	His		OK
2	MED4	pET21b	Nde I, Not I	His		OK
3	MED21	pET24d GstTEV	Sal I, Not I	His	bicistron I	2nd RBS in ORF
4	MED21	pET24d GstTEV	Sal I, Not I	His	bicistron I	2nd RBS mutated

5	MED4 MED21	pET24d GstTEV	Nhe I, EcoR I Sal I, Not I	Gst His	bicistron I	2nd RBS in ORF
6	MED10 MED21	pET24d GstTEV	Nhe I, EcoR I Sal I, Not I	Gst His	bicistron I	2nd RBS in ORF
7	MED7 MED21	pET24d GstTEV	Nhe I, EcoR I Sal I, Not I	Gst His	bicistron I	2nd RBS in ORF
8	MED6 MED21	pET24d GstTEV	Nhe I, EcoR I Sal I, Not I	Gst His	bicistron I not sequenced	2nd RBS in ORF
9	MED10 MED21	pET24d GstTEV	Nhe I, EcoR I Sal I, Not I	Gst His	bicistron I	none
10	MED4 MED21	pET24d GstTEV	Nhe I, EcoR I Sal I, Not I	Gst His	bicistron I	none
11	MED7 MED21	pET24d GstTEV	Nhe I, EcoR I Sal I, Not I	Gst His	bicistron I	none
12	MED6 MED21	pET24d GstTEV	Nhe I, EcoR I Sal I, Not I	Gst His	bicistron I not sequenced	none
13	MED10 MED21	pET21b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	good
14	MED4 MED21	pET21b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	good
15	MED7 MED21	pET21b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	good
16	MED6 MED21	pET21b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	good
17	MED7 MED4	pET21b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	good
18	MED7 MED4	pET24d GstTEV	Nhe I, EcoR I Sal I, Not I	Gst His	bicistron I	expressed
19	spMED7	pET15b	Nde I, BamH I	His	Clas Gustaffson	Thrombinsite good
20	spMED10	pET15b	Nde I, BamH I	His	Clas Gustaffson	Thrombinsite good

21	spMED4	pET15b	NdeI, BamH I	His	Clas Gustaffson	Thrombinsite
22	spCDK8	pET15b	NdeI, BamH I	His	Clas Gustaffson	Thrombinsite
23	spMED10 scMED21	pET21b	Nde I, BamH I Sal I, Not I	none His	chimera bicistron II	none
26	spMED4 scMED21	pET21b	Nde I, BamH I Sal I, Not I	none His	chimera bicistron II	none
28	spMED7 scMED21	pET21b	Nde I, BamH I Sal I, Not I	none His	chimera bicistron II	good
30	MED4ΔC731	pET21b	Nde I, Not I	His	mutation	unstable
32	MED4ΔC625	pET21b	Nde I, Not I	His	mutation	expressed
33	MED4ΔC625 MED21	pET21b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	unstable com- plex
34	MED4ΔC731 MED21	pET21b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	unstable com- plex
35	spMED21	pET24d GstTEV	Nde I, Not I	His	bicistron I	
36	MED21ΔC37	pET24d GstTEV	Nde I, Not I	His	bicistron I	none
37	MED10 MED21ΔC37	pET21b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	none
38	MED4ΔC625 MED21ΔC37	pET21b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	none
39	MED6 MED21ΔC37	pET21b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	none
40	MED7 MED21ΔC37	pET21b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	none
41	spMED10 spMED21	pET21b	Nde I, BamH I Sal I, Not I	none His	bicistron II	none
42	spMED4 spMED21	pET21b	Nde I, BamH I Sal I, Not I	none His	bicistron II	none

43	spMED7 spMED21	pET21b	Nde I, BamH I Sal I, Not I	none His	bicistron II	good
44	MED4 MED21 Δ C37	pET21b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	none
45	MED7 Δ N102 MED21	pET21b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	very good
46	MED10 MED7 MED21	pET21b	Nhe I, EcoR I EcoR I, Sal I Sal I, Not I	none none His	tricistron I	expressed
47	MED7	pET21b	Nde I, Not I	His	Claudia Buchen	expressed
48	MED7 Δ N102 MED21	pET24b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	very good
49	MED6 MED21	pET24b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	expressed
50	MED7 MED21 Δ C104	pET24d GstTEV	Nhe I, EcoR I Sal I, Not I	Gst His	bicistron I	n.a.
51	MED7 MED21 Δ N37	pET24d GstTEV	Nhe I, EcoR I Sal I, Not I	Gst His	bicistron I	n.a.
52	MED7 MED21 Δ N37 Δ C104	pET24d GstTEV	Nhe I, EcoR I Sal I, Not I	Gst His	bicistron I	n.a.
53	MED7 MED21 Δ C104	pET21b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	expressed
54	MED7 MED21 Δ N37	pET21b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	expressed
55	MED7 MED21 Δ N37 Δ C104	pET21b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	expressed
56	MED7 Δ N121 Δ C205 MED21	pET21b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	expressed
57	MED7 Δ N102 Δ C207 MED21	pET21b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	expressed
58	MED7 Δ N102, MED21	pET24b	Nhe I, EcoR I Sal I, Xho I	none none	bicistron II	very good

59	MED10 Δ N9	pET21b	Nde I, EcoR I	none		weak
60	MED10	pET21b	Nde I, EcoR I	none		expressed
61	MED7 Δ N121 MED21	pET21b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	expressed
62	MED7 Δ N102 MED21 Δ C104	pET21b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	expressed
63	MED7 Δ N102 MED21 Δ N37 Δ C103	pET21b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	expressed
64	MED7 Δ N102 Δ C207 MED21	pET24b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	expressed
65	hACID1 393-548	pET21b	Nde I, Xho I	His	Thomas Uhlmann	very good
66	MED7 Δ N102 Δ C207 MED21	pET24b	Nhe I, EcoR I Sal I, Xho I	none none	bicistron II	very good
67	MED7 Δ N102 Δ C205 MED21 Δ N37	pET24b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	very good
68	MED7 Δ N102 Δ C205 MED21 Δ N37 Δ C103	pET24b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	very good
69	MED7 Δ N102 Δ C205 MED21 Δ C103	pET24b	Nhe I, EcoR I Sal I, Not I	none His	bicistron II	very good
70	MED7 Δ N102 Δ C205 MED21 Δ C132	pET21b	Nhe I, EcoR I Sal I, Not I	none none	MED21 mutation: L5M/L119M/ L125M bicistron II	very good
71	hACID1393-548	pET24b	Nde I, Not I	none		very good
72	MED7 Δ N102 Δ C205 MED21 Δ C132	pET21b	Nhe I, EcoR I Sal I, Not I	none none	MED21 mutant: L5M/L119M/ L125M bicistron II	very strong
73	MED9	pET21b	Nde I, BamH I	none	BamH I in coding sequence	n.a.

74	VP16	pGFxcT01		Gst	Michael Meisterernst	weak
75	MED9 MED4	pET21b	Nde I, BamH I Sal I, Not I	none His	bicistron II	very strong
76	MED1	pET24d	Nco I, Not I	His		not detectable
77	MED10	pET24b	Nde I, Not I	His		OK
78	hACID1 393-548	pET24b	Nde I, Xho I	His	Thomas Uhlmann	good
79	VP61H1	pGFxcT01	BamH I, EcoR I	Gst		expressed
80	MED6 Δ C1-214	pET11a		His	Claudia Buchen	weak
81	VP61H1Mut	pGFxcT01	BamH I, EcoR I	Gst	Michael Meisterernst	expressed
82	MED7 Δ N102 Δ C205 MED21	pET24b	Nhe I, EcoR I Sal I, Not I	none none	MED7 mutant: L34M/L40M/ bicistron II	n.a.
83	MED6	pET11a		His- Flag	Claudia Buchen	weak
84	MED21	pET24b	Nde I, Not	none		expressed
85	MED9 Δ C134 MED4	pET21b	Nde I, BamH I Nde I, Not I	none His		not stoichio- metric
86	MED7 Δ N102 Δ C205 MED21 Δ C132	pET24b	Nhe I, EcoR I Sal I, Not I	none none	MED21 mutant: L5M/L119M/ L125M bicistron II	very strong
87	MED6	pGEX3X		His	88 without MED17	weak
88	MED17 Δ N241-688 MED6	pGEX3X		Gst His	Sabine Hoeppner Xa site	weak
89	dACID1 514-687	pET24b	Nde I, Xho I	His		very strong

90	dACID2 514-674	pET24b	Nde I, Xho I	His		very strong
91	MED4	pET21b	Nde I, Xho I	none		very strong
93	dACID3 514-656	pET24b	Nde I, Xho I	His		good
94	dACID4 514-663	pET24b	Nde I, Xho I	His		very strong
95	MED31	pET24d	Nco I, Not I	none		weak
96	MED31 MED10	pET24d	Nco I, EcoR I Sal I, Not I	none His	bicistron III	good
97	CTD linker (1437-1773)	pET24d GstTEV	Nhe I, BamH I	Gst	TEV site	weak
98	CTD (1556-1773)	pET24d GstTEV	Nhe I, BamH I	Gst	TEV site	weak
99	MED9 MED4	pET21b	Nde I, BamH I Sal I, Xho I	none none	bicistron II	very strong
100	MED31 MED9 MED4	pET24d	Nco I, EcoR I EcoR I, Sal I Sal I, Xho I	His none none	tricistron II	n.a.
101	MED31 MED9	pET24d	Nco I, EcoR I EcoR I, Sal I	His none	bicistron III	n.a.
102	MED31	pET24d	Nco I, EcoR I	His		expressed
103	MED31ΔN	pET24d	Nco I, EcoR I	His		weak
104	MED10 MED7 MED21	pET21b	Nhe I, EcoR I Nco I, Sal I Sal I, Xho I	none none none	tricistron I	good
105	hACID2 394-543	pET21b	Nde I, Not I	His		good
106	hACID2 394-543	pET21b	Nde I, Not I	none		n.a.
107	hACID2 394-543	pET21b	Nde I, Not I	His	K411/413A D418A	good
108	hACID2 394-543	pET21b	Nde I, Not I	none	K411/413A	n.a.
109	hACID2 394-543	pET21b	Nde I, Not I	His	E437/442A K440A	good
110	hACID2 394-543	pET21b	Nde I, Not I	none	E437/442A K440A	n.a.
111	hACID2 394-543	pET21b	Nde I, Not I	His	K478A	good

112	hACID2 394-543	pET21b	Nde I, Not I	none	K478A	n.a.
113	hACID2 394-543	pET21b	Nde I, Not I	His	K518/519/ 520A	good
114	hACID2 394-543	pET21b	Nde I, Not I	none	K518/519/ 520A	n.a.

Table 8: Plasmids

4.1.3 Growth media

Media	Reference	composition
LB	(Miller, 1972)	1 % tryptone; 0.5% yeast extract; 0.5% NaCl
Minimal Media	Budisa et al., 1995; Meinhart et al., 2003a	7.5 mM ammonium sulfate, 8.5 mM NaCl, 55 mM KH_2PO_4 , 100 mM K_2HPO_4 , 1mM MgSO_4 , 20mM glucose, 1 $\mu\text{g/l}$ Trace elements (Cu^{2+} , Mn^{2+} , Zn^{2+} , Mo_4^{2-}), 10 mg/l Thiamine, 10 mg/l Biotine, 1 mg/l Ca^{2+} , 1 mg/l Fe^{2+} , 100 mg/l amino acids (A, C, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, W, Y) , 100 mg/l selenomethionine.

Table 9: Growth media

supplements	stock solution	in Media
ampicilin	100 mg/ml in H_2O	100 mg/l
kanamycin	50 mg/ml in H_2O	50 mg/l
chloramphenicol	50 mg/ml in ethanol	50 mg/l
IPTG	1 M in H_2O	0.5 mM

Table 10: Supplements

4.1.4 Bioinformatic tools

Homology searches and alignments Homologous sequences were found using the NCBI BLAST and PSI BLAST-Server: <http://www.ncbi.nlm.nih.gov/BLAST/> or from the publication of Boube et al. (Boube et al., 2002). Multiple sequence alignments were done with ClustalW of the program MaC Vector (accelrys) using default settings and displayed by the programs Amas and Alscript (Barton, 1993).

Secondary structure prediction Secondary structure prediction was done using the PredictProtein (<http://cubic/bioc/columbia/edu/predictprotein.html>) secondary structure prediction server. Generally a multiple sequence alignment was sent, since the accuracy of secondary structure prediction depends on the quality of the alignment. For prediction of coiled-coil regions in the proteins the program coils (www.ch.embnet.org/software/COILS_form.html) was used.

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

4.1.5 Protein expression and selenomethionine labeling

Proteins were routinely expressed recombinantly in *E. coli*. Plasmids containing the desired protein variants were transformed into competent *E. coli* BL21 DE3 (Stratagene). Cells were grown at 37° C in LB medium supplemented with the antibiotic corresponding to the resistance cassette of the plasmid. Once the cells grew to an OD₆₀₀ of 0.5 cultures were cooled on ice for 30 min. Protein expression was induced by addition of 0.5 mM IPTG and cells were grown at 20° C over night. Cells were harvested by centrifugation (5000 rpm, SLS6000 rotor) at 4° C, resuspended in lysis buffer (see corresponding purification) and flash frozen in liquid nitrogen. Cell pellets were stored at -80° C.

For selenomethionine labeling of MED7 Δ N Δ C/MED21, MED7 Δ N Δ C/MED21D Δ C, MED7 Δ N Δ C/MED10His/MED21His and MED4His/MED7 Δ N Δ C/MED21His plas-

mid DNA containing the corresponding genes were transformed into the methionine auxotroph *E.coli* strain B834 (DE3) (Stratagene). Bacteria were grown in LB medium supplemented with the appropriate antibiotic at 37° C to an OD₆₀₀ of 0.5. Cells were harvested and resuspended in the same amount of minimal medium (Table 9) supplemented with selenomethionine (100 mg/l) and antibiotics. Cells were grown until the OD₆₀₀ increased by 0.2 at 37° C to deplete the medium of any residual methionine. Cultures were cooled on ice for 30 min and protein expression was induced by the addition of 0.5 mM IPTG. Protein was expressed over night at 18° C. Incorporation of selenomethionine was confirmed by mass spectrometry of the purified proteins.

4.1.6 Limited proteolysis experiments

For chymotrypsin and trypsin treatment 1 μ g of the corresponding protease was added to 20 μ g to 50 μ g of purified protein. Digests were done in the buffers used for gelfiltration and supplemented with CaCl₂ to a final concentration of 4 μ M. The mixture was incubated at 37° C for 1 min, 3 min, 10 min, 30 min, and 60 min. The reactions were stopped by the addition of SDS sample buffer and were heated immediately to 95° C for 5 min.

For proteinase K treatment 1 μ l of dilutions of proteinase K (3 μ g/ μ l, 0.3 μ g/ μ l, 0.03 μ g/ μ l, 0.003 μ g/ μ l) were added to the protein samples. The mixtures were incubated on ice for 1 h. The reactions were stopped by the addition of sample buffer and boiling as above. All samples were analyzed by SDS-PAGE. Bands of interest were passively transferred to PVDF membrane and analyzed by Edman sequencing as described in 4.1.7.

4.1.7 Protein Analysis

Protein separation by SDS-PAGE For separation of protein samples glycine-SDS-PAGE (10 % - 20 % acrylamide; Laemmli, 1970) was performed. Gels were then either subjected to protein transfer for Western-blot or directly stained with Coomassie (SIGMA) solution. Silver staining (Bloom et al., 1987) was performed when very low amounts of proteins had to be visualized.

Passive adsorption of electrophoresis samples onto PVDF Membranes and Edman sequencing For N-terminal sequencing proteins were separated on SDS-PAGE and stained with coomassie. The protein band of interest was excised and dried in a speed vac. After drying the piece of gel was rehydrated in 20 ml of 200 mM Tris pH 8.5 (room temperature), 2 % SDS. Then 100 μ l of distilled water was added to set up a concentration gradient together with a small piece of pre wet (ethanol) PVDF membrane. Once the solution turned blue 10 μ l of methanol was added as a catalyst. After 1 to 2 days the solution became clear and the membrane piece blue, indicating a complete transfer. The membrane was washed 5 times with 10 % Methanol vortexing 30 sec each time. The Protein was N-terminally sequenced from the dry membrane in a PROCISE 491 sequencer (Applied Biosystems).

4.1.8 Gelfiltration experiments and calculation of the hydrodynamic radius from the structure

To estimate the molecular weight of proteins and complexes, analytical gelfiltration chromatography was performed in the corresponding protein buffers. A Superose-6 HR gelfiltration column (Amersham) was used for these experiments. The column was calibrated with the standard proteins thyroglobulin (669 kDa), ferritin (440 kDa), catalase (232 kDa) and aldolase (158 kDa) (Amersham) immediately before the experiment. To obtain an estimate of the stoichiometry of the MED7/MED21 complex the hydrodynamic radii of monomers and dimers were compared to those of the standards. The hydrodynamic radii of all proteins were calculated with the program HYDROPRO (Garcia De La Torre et al., 2000) from the structures of the corresponding proteins.

4.1.9 Static light scattering

Static light scattering measurements were performed with a triple detector TDA (viscotek) connected to a Superose 6 gelfiltration column equilibrated with 150 mM KCl, 10 mM Hepes pH 7.3, 5 mM DTT. 200 μ l protein with a concentration of 1.8 mg/ml was injected onto the column. The concentration/volume was calculated from the UV absorption and extinction coefficient. With this information the refractive index/volume was calculated by means of the refractive index detector. Refractive index, UV and viscosity were followed during the measurement. The hydrodynamic radius

and the molecular weight were calculated using the static light scattering software from viscotek with the calculated refractive index.

4.2 The middle module

4.2.1 Strategy for the expression of the middle module proteins

Most of the middle module proteins are only very poorly soluble when expressed alone. This problem was overcome using multicistronic expression, where two or more open reading frames are under the control of one single promoter. During transcription the open reading frames are transcribed onto one single mRNA and thus translation of all proteins can occur close in time and space. This allows the nascent polypeptides to associate with their partners and fold correctly. To coexpress proteins on different vectors cells were cotransformed with plasmids containing the additional protein and a different antibiotic resistance cassette .

4.2.2 Cloning

Proteins of the middle module were cloned into pET 21b or pET 24b (Novagen) with or without C-terminal hexahistidine tag (His) according to table 8. For bicistronic expression of different middle module proteins together with MED21His the genes of the corresponding proteins were cloned into the pET21b (Novagen) or pET 24b (Novagen) vector using the restriction sites Nhe I and EcoR I. Having Nhe I as the N-terminal restriction site introduces three additional amino acids at the N-terminus. The generated N-terminus is MAS. The methionine is cleaved off in *E.coli* and leaves the N-terminal AS free for Edman sequencing. The gene of MED21 was then inserted, together with a second ribosomal binding site, at the 5' region of the MED21 gene as described (Lutzmann et al., 2002). The design of the primer used can be seen in Figure 28. The restriction sites used were Sal I and Not I. An additional Nde I site is inserted at the start site of the MED21 ORF for easier shuffling of the genes. Bicistronic expression of MED4/MED9 and the tricistrones of MED4/MED9/MED31 and MED7/MED10/MED31 were done in the same way. The exact restriction sites can be seen in table 8. All *S. cerevisiae* proteins were amplified from *S. cerevisiae* genomic DNA. Constructs encoding for *S. pombe* proteins were done using *S. pombe* cDNA obtained from C. Gustafsson.

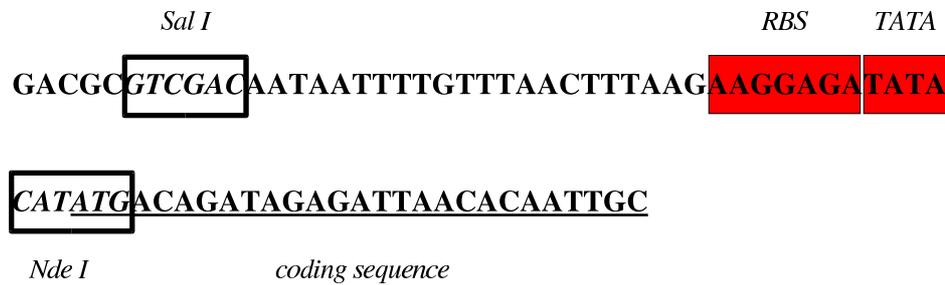


Figure 28: The forward primer used for bicistronic expression: it contains an additional ribosomal binding site (RBS), The coding sequence of MED21 is underlined.

Point mutations Point mutations were generated to introduce a silent mutation in the MED21 ORF to remove an internal ribosomal binding site (RBS). MED21 was mutated to contain additional methionines for structure solution. All point mutations were introduced by site-directed mutagenesis with the two step PCR-method. The gene of interest was amplified from the cloned plasmid in two steps. In the first step an oligonucleotide containing the mutations was used as a reverse primer together with the forward primer of the gene of interest. In the second step the generated PCR product -which contains the mutations- was used instead of the forward primer together with a reverse primer of the gene. The resulting PCR product was digested and ligated into the corresponding vector.

4.2.3 Expression and purification of middle module proteins

MED7/MED21 variants Plasmids containing different variants of MED7/MED21 were transformed into competent BL21 DE3 (Stratagene) and expressed as described in 4.1.5. For protein purification cells were thawed at 30° C and lysed by 15 min of sonication using a flat $\frac{1}{2}$ " working tip with 20 % duty time and 40 % output on a Branson sonifier. The slurry was cleared by 30 min of centrifugation in a SS34 rotor at 4° C and 15000 rpm. The complexes containing a His-tag were loaded onto a 3 ml self assembled Ni-NTA agarose column (Qiagen) equilibrated with lysis buffer. After washing with 30 ml lysis buffer bound protein was eluted with 15 ml lysis buffer containing 200 mM imidazole. Eluted proteins were diluted 1:1 with MonoQ buffer containing

50 mM NaCl and further purified by anion exchange chromatography (MonoQ 10/100 GL (Amersham)). After binding the column was washed with two column volumes (CV) of MonoQ buffer containing 50 mM NaCl. Bound proteins were eluted with a 10 CV linear gradient of the same buffer containing 50 mM to 1 M NaCl. After concentration, the pooled fractions were applied to a Superose-6 HR gel filtration column (Amersham) equilibrated with MED7/MED21 buffer. Peak fractions were analyzed by SDS-PAGE, combined and concentrated to a final concentration of 16 mg/ml.

For purification of the MED7/MED21 variants without an affinity tag proteins in the cleared lysate were precipitated by the addition of saturated ammonium sulfate solution to a final concentration of 35 %. After 30 min of centrifugation at 4° C and 12 000 rpm in a SS34 rotor the pellet was dissolved in MonoQ buffer containing 50mM NaCl. Buffer was added until the conductivity was below 100 mS/cm³ and the sample was further purified in the same way as the affinity purified MED7/MED21 complex.

Purification of tri and tetrameric complexes A plasmid containing the MED7 Δ N Δ C/MED21His and either MED4His or MED10His were cotransformed into *E. coli* BL21 DE3 (Stratagene) and expressed as described (4.1.5). The tetrameric MED4/MED7/MED9/MED21 complex was obtained by cotransforming a plasmids containing MED4/MED9 and MED7 Δ N Δ C/MED21His. For all three complexes the purification procedure of MED7 /MED21 variants was followed until the MonoQ where elution was done using 20 CV linear gradient and dimeric MED7/MED21 complexes could be separated from trimeric MED4/MED7/MED21 or MED7/MED10/MED21. Further polishing of the samples was done on a Superose 6 gel filtration column in MED7/MED21 buffer as above.

Purification of MED4/MED9 MED4 and MED9 are coexpressed from one single plasmid in a bicistronic manner. None of the two proteins has an affinity tag. For purification the cleared lysate was precipitated by the drop wise addition of saturated ammonium sulfate solution to a final concentration of 25 %. The mixture was stirred on ice for 30 min and then centrifuged in a SS34 rotor for 30 min at 4° C and 12000 rpm. The pellet was dissolved in lysis buffer and remaining debris was removed by an other centrifugation step (20 min 16000 rpm at 4° C). The protein concentration was determined by Bradford reagent (Biorad).

Name	Description
lysis buffer	150 mM NaCl, 20 mM Tris pH 8.5 (4° C), 10mM β -mercaptoethanol, 1 mM PMSF, 0 to 200mM imidazole
MonoQ buffer ammonium sulfate	50 mM to 1 M NaCl, 20 mM Tris pH 8.5 (4° C), 3mM DTT at room temperature saturated solution of ammonium sulfate
Gst buffer	150 mM NaCl, 20 mM Tris pH 7.5 (4° C), 10mM DTT, 0 to 100 mM reduced glutathione
MED7/MED21 buffer	150mM NaCl, 20 mM Bicine pH 8.5 (4° C), 3 mM DTT

Table 11: Buffers used during the purification of MED7/MED21 variants

Purification of MED7/MED4 A plasmid containing GstMED7 and MED4His was expressed, cells were harvested and cleared lysate was prepared as described above. The protein complex was first purified on a Ni NTA sepharose (Qiagen) gravity flow column due to the His tag on MED4. The eluate of the Ni NTA was adjusted to pH 7.5 by dilution with lysis buffer containing 20 mM Tris pH 7.5 (4° C) and 10 mM DTT. 500 μ l of Gst beads (CL4B, Amersham) were added to the sample and the mixture was incubated on a rotating wheel for 4 h at 4° C. After three washing steps bound protein was eluted with 100 mM reduced glutathione in lysis buffer.

MED31 purification MED31 containing a His tag was expressed as described above. Purification was as for MED7/MED21His except for the Ni affinity step where a 1 ml HisTrap HP (Amersham) was used. The cleared lysate was loaded onto the preequilibrated (lysis buffer) HisTrap column on a peristaltic pump at 0.5 ml/min. The column was washed with Lysis buffer. Bound protein was eluted on a Aekta FPLC system (Amersham) with a stepwise gradient of 20, 40, and 200 mM imidazole. Elution of protein was monitored by absorption at 280 nm. Purity of the protein was analyzed by SDS-PAGE. Further purification was done by anion exchange chromatography and gelfiltration as above.

4.2.4 Interaction assay

To test interactions of MED31 with other subunits of the middle module a plasmid containing the gene of MED31 was cotransformed with a plasmid coding for one of

the following proteins: MED4, MED6, MED7, MED9, MED10 and MED21. Each of these proteins was cloned such that it contains a his tag except for MED9 which was cotransformed with MED31His. Protein expression was done as described above, in small cultures of 50 ml. Cells were harvested by 15 minutes of centrifugation at 4500 rpm in 50 ml falcon tubes in a Haereus centrifuge. Cells were resuspended in 2 ml lysis buffer. For lysis glass beads were added to the suspension and cells were opened by vigorous shaking for 30 minutes on a home made shaking machine. Non soluble debris was removed by 3 minutes centrifugation in a micro centrifuge at 14000 rpm. The supernatants were incubated with 30 μ l of prewashed Ni NTA matrix for 3h. Beads were pelleted by centrifugation (700 g, in a micro centrifuge) and washed 4 times with lysis buffer. All steps were performed at 4° C. Bound proteins were eluted by the addition of SDS sample buffer, boiled for 5 min at 95° C and analyzed by SDS-PAGE.

4.2.5 Crystallization and crystal treatment

Crystallization Variants of MED7/MED21 were crystallized at 20° C with the hanging drop vapor diffusion method using 24 well plates and plastic cover slides (molecular dimensions). An initial crystallization condition of MED7 Δ N Δ C/Med21 stems from the commercial Hampton crystal screens II and contains 10 % ethanol and 1.5 M NaCl. This condition could be refined as mentioned in table 11. Mutant MED7 Δ N Δ C /MED21 Δ C crystallized in additional 4 conditions in the Hampton screen I (0.4 M NaK tartrate; 30 % PEG400, 0.1 M Tris pH 8.5, 0.2 M Na citrate; 1 M Na acetate, 0.1 M imidazole pH 6.5; 0.2M Mg formate) and could be optimized as shown in table 11. Crystals of higher order complexes of MED10/MED7 Δ N Δ C/MED21 and MED4/MED7 Δ N Δ C/MED21 were obtained in very similar conditions.

Crystal freezing Crystals obtained in acetate and tartrate were harvested in mother solution, which was exchanged gradually against mother solution containing additional 35 % glycerol. The gradual exchange of the solutions was done in then steps were in a first step 1/10 of the drop solution was replaced by cryosolution, in a second step 1/9 and so on. This steps were done until replacement of 1/2 of the solution, which was repeated twice to ensure complete replacement of the mothersolution by the cryosolution. Crystals were flash frozen in liquid nitrogen.

Crystals of the native and selenomethionine MED7 Δ N Δ C/MED21His complex were grown in mother solutions containing ethanol. For freezing of crystals grown in these conditions the mother solution was exchanged by the same solution containing 25 % glycerol, but no ethanol, using microdialysis buttons.

4.2.6 X-ray structure determination

Data collection: All diffraction data were collected at the beamline X06SA at the Swiss Light Source, Villigen, Switzerland. The crystals were radiation sensitive and diffraction power was rather weak. Therefore images were recorded with maximal possible detector distance to reduce the background. To further reduce the background images were collected with an increment of 0.5°/image. Focusing of the beam on the detector rather than on the crystal helped to increase the resolution and decrease radiation damage of the crystal.

For MAD phasing, three additional methionines were introduced at positions of conserved hydrophobic residues in MED21 Δ C (L5M/L119M/L125M). MAD experiments were performed on crystals of selenomethionine-labeled protein, and diffraction data were processed with DENZO and SCALEPACK (Ottwinowski and Minor, 1996) (Table 3).

Structure solution Four selenium sites were correctly identified with program SOLVE (Terwilliger, 2002), and refined using SHARP (La Fortelle and Bricogne). Three additional selenium peaks were detected in residual electron density maps with SHARP. Phasing with SHARP, using seven consistent peaks, led to an electron density map that revealed distinct α -helices. A total of ten methionine residues are present in the amino acid sequences and were ordered in the electron density map, except the two N-terminal methionines and M42 of MED21. Using the selenium sites as sequence markers, most of MED7 and MED21 could be built into the electron density map at 3.6 Å resolution with program O (Jones et al., 1991). This initial model was repositioned in the unit cell of the native crystal by rigid body refinement with CNS (Brunger et al., 1998). Model-phased maps at the final resolution of 3.0 Å allowed us to complete and refine the structure with CNS to a free R-factor of 28.9 % (Table 4). The second crystal form, belonging to space group C222₁ was phased using the 4-helix bundle of the refined MED7/MED21 structure. A difference electron density map phased with the

Variant	Crystallization condition	Cryo-condition	used for
MED7 Δ N Δ C MED21His	0.8-1.1 M NaCl, 10-15 % ethanol, 0-3 % PEG 6000, 10 mM β -mercaptoethanol, 10 mM EDTA	0.8-1.1 M NaCl, 25 % glycerol 0-3 % PEG 6000, 10 mM Bicine pH 8.5, 10 mM β -mercaptoethanol, 10 mM EDTA	molecular replacement, C222 ₁
MED7 Δ N Δ C MED21 Δ C	500 mM sodium acetate, 100 mM Hepes pH 7.5, 5 % PEG 400, 10 mM β -mercaptoethanol	crystallization condition containing additionally 35 % glycerol	native dataset, phase extension, P4 ₃ 22
MED7 Δ N Δ C MED21 Δ C	100 mM sodium potassium tartrate, 100 mM Hepes pH 7.5, 5 % PEG 400, 10 mM β -mercaptoethanol, 10 mM EDTA	crystallization condition containing additionally 35 % glycerol	selenomethionine derivative, P4 ₃ 22
MED7 Δ N Δ C MED10 MED21	0.8 M NaCl, 11 % ethanol, 10 mM β -mercaptoethanol, 10 mM EDTA	0.8 M NaCl, 30 % glycerol, 10 mM Bicine pH 8.5, 10 mM β -mercaptoethanol, 10 mM EDTA	molecular replacement, C222 ₁
MED7 Δ N Δ C MED10 MED21	160 mM sodium potassium tartrate, 10 % ethanol, 10 mM β -mercaptoethanol, 10 mM EDTA	160 mM sodium potassium tartrate, 35 % Glycerol, 10 mM Bicine pH 8.5, 10 mM β -mercaptoethanol, 10 mM EDTA	molecular replacement, C222 ₁
MED4 MED7 Δ N Δ C MED21	300mM ammonium acetate, 3 % Isopropanol, 10 mM 10 mM β -mercaptoethanol, 10 mM EDTA	300mM ammonium acetate, 5 % Glycerol, 10 mM Bicine pH 8.5, 10 mM β -mercaptoethanol, 10 mM EDTA	Molecular replacement, C222 ₁
MED4 MED7 Δ N Δ C MED21	0.3-0.7 M NaCl, 7 % ethanol, 10 mM β -mercaptoethanol, 10 mM EDTA	0.3-0.7 M NaCl, 35 % glycerol, 10 mM Bicine pH 8.5, 10 mM β -mercaptoethanol, 10 mM EDTA	selenomethionine derivative, C222 ₁

Table 12: Crystallization solutions of variants of MED7/MED21, MED4/MED7/MED21 and MED7/MED10/MED21

bundle domain showed locations of the coiled-coil protrusions in the two heterodimers of the asymmetric unit that deviated substantially from that observed in the original structure. The coiled-coil region was adjusted. An extended loop between MED21 helices $\alpha 1$ and $\alpha 2$ that was disordered in the P4₃2₂ crystal form was added as poly-alanine in one of two heterodimers in the asymmetric unit, and the resulting model was refined to a free R-factor of 31.4 % (Table 4). The MED21 loop $\alpha 1$ - $\alpha 2$ is better ordered in the C222₁ crystal form since it forms a crystal contact that is not observed in the P4₃2₂ crystal form. In both refined structures, none of the residues fall in disallowed regions of the Ramachandran plot (Table 4).

4.3 Reconstitution of the middle module

4.3.1 Expression and purification of the MED7/MED10/MED21/MED30 middle module complex

Plasmids containing MED7/MED10/MED21 and MED31 were cotransformed into BL21 DE3 RIL and expressed as described in 4.1.5. For purification of expressed proteins cells were thawed at 30° C and lysed by sonication with a Branson sonifier 250 and a flat $\frac{1}{2}$ " working tip for 15 min 20 % duty time and 40 % output. The lysate was then centrifuged for 30 min at 4° C and 16000 rpm in a SS34 rotor. The cleared lysate was loaded onto a preequilibrated 1 ml HisTrap HP (Amersham) column washed with 20 mM of lysis buffer and eluted stepwise with lysis buffer containing 20, 40, 70 and 200 mM imidazole. Protein elution was monitored by absorption at 280 nm and Bradford reagent (Biorad). Eluted fractions were analyzed by 17 % SDS-PAGE. The main fractions of the 70 mM imidazole elution were diluted with buffer A and loaded on a MonoQ 10/100 GL (Amersham) anion exchange chromatography column and eluted with a 20 column volume linear gradient of buffer A containing 50 mM to 1 M NaCl. The main fractions were analyzed by 17 % SDS-PAGE, pooled and concentrated in Amicon Ultra centrifugal devices with 10 KDa molecular weight cut off. The concentrated samples were loaded onto a Superose 6 10/300 GL (Amersham) gelfiltration column equilibrated with Buffer B. Peak fractions were analyzed by 17 % SDS-PAGE, concentrated to 4.5 mg/ml and used for crystallization setups.

4.3.2 Reconstitution of a MED4/MED7/MED9/MED10/MED21/MED30 middle module complex

The 40 mM His Trap elution of MED7/MED10/MED21/MED30 was mixed in a 1 to 1.5 ratio with ammonium sulfate purified MED4/MED9 (4.2.3) and the complex was assembled for 1 h at 20° C. After 1h the conductivity was adjusted to less than 100 mS/cm³ by the addition of buffer A and the sample was loaded on a MonoQ 10/100 GL (Amersham). Elution was as for MED7/MED10/MED21/MED30. Pure fractions were concentrated in Amicon Ultra centrifugal devices (10 KDa molecular weight cut off) and applied to a Superose 6 10/300 GL (Amersham) gelfiltration column preequilibrated with buffer B. Peak fractions were analyzed by SDS-PAGE, concentrated to a final concentration of 5 mg/ml and used for Crystallization setups as well as for Pol II middle module complex assembly.

4.3.3 Assembly of a Pol II/middle module complex

For reconstitution of a middle module Pol II complex 10 subunit Pol II, Rpb4/Rpb7 and the purified middle module complex are assembled. Pol II and Rpb4/Rpb7 were obtained from Stefan Benkert. Buffers of 10 subunit Pol II and middle module were exchanged to Pol II buffer using Amicon Ultrafree-MC centrifugal devices with 100 and 10 KDa cut off, respectively. Rpb4/Rpb7 were obtained frozen in Pol II buffer and thawed on ice. 10 subunit Pol II, Rpb4/Rpb7, and middle module were combined and incubated on a rotating wheel at 20° C for 1 h prior to gelfiltration on Superose 6 10/300 GL (Amersham). 5 fold excess of Rpb4/Rpb7 and 3 fold Excess of middle module were used the total volume was adjusted to 200 μ l. Peak fractions were analyzed by 17 % SDS-PAGE, fractions containing the complex were concentrated to 200 μ l and reloaded onto a Superose 6 gelfiltration column. The peak fractions were again analyzed by 17 % SDS-PAGE.

buffer	description
lysis buffer	150 mM NaCl, 20 mM Tris pH 8.5 (4° C), 10 mM β -mercaptoethanol, 1 mM PMSF, 0 to 200 mM imidazole
MonoQ buffer	50 mM to 1 M NaCl, 20 mM Tris pH 8.5 (4° C), 10 mM β -mercaptoethanol
ammonium sulfate	at room temperature saturated solution of ammonium sulfate
middle module buffer	150 mM NaCl, 20 mM Tris pH 8.5 (4° C), 10 mM β -mercaptoethanol
Pol II buffer	40 mM ammonium sulfate, 5 mM Hepes pH 7.25 10 μ M ZnCl ₂ , 10 mM DTT

Table 13: Buffers used for the purification and reconstitution of the middle module

4.3.4 Gst-CTD pull-down

Gst-CTD pull-downs were done to test binding of the six-subunit middle module complex to the CTD of Pol II. Gst fusion proteins of the CTD repeats only and the CTD containing the linker domain of Rpb1 were used. Both Gst-CTD constructs as well as a Gst only control were expressed as described in 4.1.5. The cells were resuspended in 50 HGN100 buffer and lysed by sonication as described above. After two 30 minutes centrifugation steps (SS34, 16000rpm, 4° C). 500 μ l of prewashed Gst sepharose (CL4B, Amersham) was added to the cleared lysate. Gst fusion protein was bound to the beads during 4 hours at 4° C on a rotating wheel. The slurry was decanted into plastic columns (Qiagen) and washed with 20 ml of HEGN100 buffer. Amounts and purity of the bound protein was analyzed by Bradford and SDS-PAGE. Protein amounts were equalized by diluting with empty Gst sepharose to a final concentration of 0.5 μ g/ μ l beads. As a positive control the CTD interacting domain of Pcf11 (residues 1 to 140) was expressed and purified according to (Meinhart and Cramer, 2004). Pcf11 and the reconstituted middle module was dialyzed against HEGN100 buffer and incubated with 30 μ l Gst-CTD beads over night. The beads were washed for four times with HEGN100, boiled in SDS sample buffer and analyzed by SDS-PAGE and western blot. The HisProbe system (pierce) was used to detect His-tagged proteins.

buffer	description
HEGN100	10% glycerol, 100 μ M EDTA, 20 mM Hepes pH 7.6, 0.1 % NP40, 100 mM KCl, 1mM PMSF, 5mM DTT
BSA	10 mg/ml solution (Biolabs)

Table 14: Buffers used for the Gst-CTD pull-down

4.4 ACID

4.4.1 Design of ACID variants

Since the fulllength ACID is a large protein of 752 amino acids containing two separate domains connected by a large unordered and strongly charged loop and C-terminus, it is not suited for crystallization. The approach is therefore to determine the structure of the activator interaction domain together with the activation domain of VP16. Since crystallization of the original interaction domain did not yield crystals I used an other species *Drosophila melanogaster* and designed 4 truncations of the domain. Clusters of highly charged residues on the protein surface are often inhibiting crystallization. Therefore putative K and E surface residues were mutated to A. Design of mutations and truncations was based on results from sequence alignments, secondary structure prediction and proteolytic digests of the protein.

A first construct containing the activator interaction domain of human ACID amino acids 393 to 548 was obtained from T. Uhlmann. For cloning of the drosophila activator interaction domain of ACID; drosophila cDNA was used. Further constructs of the human and drosophila proteins were obtained by standard cloning methods using the plasmid and cDNA as templates. Mutations were introduced by the two step PCR-method using primers which contain the corresponding mutations. For cloning of Gst-VP16H1 the plasmid containing the complete VP16 activation domain (404-490) (Donaldson and Capone, 1992) was used. A premature stop codon after amino acid 451 was inserted by PCR.

4.4.2 Purification of ACID constructs

Affinity purification: Variants containing a His tag were resuspended in lysis buffer supplemented with 1 mM PMSF and lysed by French press (Gaulin) with 750 Bar

name	species	variant	size (AA)	MW	PI	ϵ	tag
hACIDHis	human	393-548	166	19.1KDa	9.6	22310	His
hACID	human	393-548	158	18KDa	9.8	22310	none
hACIDHis	human	394-543	160	18.12	9.51	22310	His
hACID	human	394-543	152	17.06	9.64	22310	none
hACIDHis	human	394-543	160	17.96	9.44	22310	His
K411/413A/D418A							
hACID	human	394-543	152	16.9	9.58	22310	none
K411/413A/D418A							
hACIDHis	human	394-543	160	17.95	9.68	22310	His
E437/442A/K440A							
hACID	human	394-543	152	16.88	9.8	22310	none
E437/442A/K440A							
hACIDHis	human	394-543	160	18.07	9.41	22310	His
K478A							
hACID	human	394-543	152	17.0	9.55	22310	none
K478A							
hACIDHis	human	394-543	160	17.95	9.1	22310	His
K518/519/520A							
hACID	human	394-543	152	16.89	9.3	22310	none
K518/519/520A							
dACIDHis 1	drosophila	515-674	172	19.6KDa	8.7	19870	His
dACIDHis 2	drosophila	515-687	185	21.2KDa	8.9	19870	His
dACIDHis 3	drosophila	515-656	154	17.4KDa	7.1	19870	His
dACIDHis 4	drosophila	515-663	161	18.3KDa	8.4	19870	His

Table 15: Different variants of drosophila and human Acid, their size, molecular weight (MW), isoelectric point (PI) and absorption coefficient ϵ at 280nm.

pressure. Lysed cells were centrifuged for 30 min at 4°C in a SS34 rotor at 16000 rpm to remove cellular debris. The supernatant was loaded twice onto a self assembled 4 ml Ni-NTA agarose (Qiagen) gravity flow column. The column was sequentially washed with 20 ml of Lysis buffer, 20 ml of Lysis buffer containing 1 M NaCl to remove bound DNA, and 20 ml of Lysis buffer containing 10 mM imidazole. Elution was done using lysis buffer containing 200 mM imidazole and a final concentration of 1 mM of EDTA was added to the eluate to chelate Ni²⁺ bleeding of the column. For the variants of the human proteins the eluate of the Ni-NTA was diluted using Source S buffer containing 50 mM NaCl until conductivity was below 100 mS/cm³ and was passed over a SourceS (Amersham) column. Bound proteins were eluted with a 10 column volume linear gradient with SourceS buffer containing 50 mM to 1 M NaCl. Variants of the Drosophila protein were passed over a MonoQ instead of the SourceS column using the corresponding buffer. The protein containing fractions were analyzed by SDS-PAGE, concentrated and loaded onto a Superdex 75 HR gel filtration column preequilibrated with hACID, dACID or NMR buffer. All fractions were analyzed by 17 % SDS-PAGE and pure protein fractions were concentrated in 10 KDa cut off Amicon Ultra centrifugal devices (millipore) to a final concentration of 10 to 16 mg/ml and used for crystallization setups. dACID 3 was concentrated to 19.3 mg/ml and used for NMR tests.

Purification of variants without an affinity tag: Variants without an affinity tag were purified by cation exchange chromatography, taking advantage of the strong positive charge of the domain as shown in table 15. Cells were resuspended in SP buffer supplemented with 5 mM DTT and 1 mM PMSF and lysed by French Press as described above. After lysis and centrifugation the lysate was loaded onto an 5 ml Hitrap SP HP (Amersham) cation exchange column. The bound proteins were eluted with a 10 column volume linear gradient of buffer SP containing 50 mM to 1 M NaCl. Protein elution was monitored by absorption at 280 nm, DNA contamination at 260 nm throughout the whole purification procedure. Protein containing fractions were analyzed by 17 % SDS-PAGE, the peak fractions pooled and diluted with MonoQ buffer as above. Care was taken that the pH reached 8.5 before loading onto a MonoQ 10/100 GL (Amersham). The protein containing flow through was concentrated and loaded onto a Superdex75 10/30 HR as described for the affinity purified ACID. The Protein containing fractions were analyzed by 17 % SDS-PAGE concentrated to a final

concentration of 10 mg/ml as above and used for crystallization setups.

4.4.3 Purification of the ACID/VP16 complex

Two plasmids containing the human ACID 393-548 and a Gst-VP16 activation domain fusion were cotransformed in BL21 (DE3) and expressed as described 4.1.5. Cell pellets were resuspended in Gst buffer containing 5 mM DTT and 1 mM PMSF lysed by French press as described above and incubated with 2 ml glutathione beads (CL4B, Amersham) for 4 h at 4° C on a rotating wheel. The suspension was decanted into an empty column, washed with 20 ml of Gst buffer, followed by a wash of 20ml of Gst buffer containing 1 M NaCl and eluted with Gst elution buffer containing 50 mM of reduced glutathione. Alternatively thrombin digestion was performed over night at 4° C using 10 U thrombin/mg of protein directly on the column. Thrombin digested proteins were further purified as above by Ni NTA using lysis buffer to dilute the sample until a pH of 8.5 was reached. Eluted fractions were concentrated and further purified by gelfiltration in hACID buffer as above. Peak fractions were concentrated as above to a final concentration of about 10 mg/ml and used for crystallization setups.

4.4.4 Binding of ACID to Gst-VP16

Binding assays were performed as described in 4.3.4 for pull downs with the Gst-CTD using 0.5 $\mu\text{g}/\mu\text{l}$ Gst-VP16 H1 and Gst-VP16 H1mutant on glutathione beads (CL4B, Amersham) and cleared *E. coli* lysates of 100 ml of overnight culture of the corresponding ACID variant.

Buffer	Description
Lysis buffer	150 mM to 2 M NaCl, 20 mM Tris-Cl pH 8.5 at 4°C, 10 mM β -mercaptoethanol, 0 to 200mM Imidazole
MonoQ buffer	50 mM to 1 M NaCl, 20 mM Tris-Cl pH 8.5 at 4°C, 5 mM DTT
SoruceS buffer	50 mM to 1 M NaCl, 20 mM Tris pH 7.0 at 4° C, 5 mM DTT
SP buffer	50 mM to 1M NaCl, 20 mM MES pH 6.5 at 4° C, 5 mM DTT
hACID buffer	150 mM NaCl, 10 mM Hepes pH 7.5, 5 mM DTT
dACID buffer	50 mM ammonium sulfate, 10 mM sodium citrate, 5 mM DTT
NMR buffer	100 mM NaCl, 20 mM Na ₂ PO ₄ pH 6.5, 10 mM β -mercapto-ethanol
Gst buffer	150 mM NaCl, 20 mM Hepes pH 7.5, 5 mM DTT
Gst elution buffer	150 mM NaCl, 20 mM Tris pH 8.0 at 4° C, 50 mM reduced glutathione, 5 mM DTT

Table 16: Buffers used during the purification of ACID and ACID VP16 variants

4.4.5 Electrophoretic mobility shift assay

The DNA binding capability of ACID was analyzed by electrophoretic mobility shift assay (EMSA). In 15 μ l of 0.5 x TBE 50 to 200 pMol ACID were incubated with 50 pMol of previously annealed DNA for 1 h at 4° C. 5 μ l of sample buffer were added and the sample was loaded on a native 8 % acrylamide gel in 0.5 x TBE buffer. Electrophoresis was carried out at 70 to 80 Volt at 4° C for 1 to 2 h. The gel was removed and stained for 20 min in SyBr gold (Molecular probes) and visualized by UV light at 310 nm. After DNA staining protein bands were stained by coomassie.

DNA:	17mer:	TTATTCCCATGGGAATA
	19mer:	TTTATTCCCATGGGAATAA
	21mer:	TTTTATTCCCATGGGAATAAA
	23mer:	TTTTTATTCCCATGGGAATAAAA
	25mer:	TTTTTTATTCCCATGGGAATAAAAA
0.5 x TBE		89 mM Tris, 89 mM boric acid, 2 mM EDTA
sample buffer		0.5 x TBE. 60 % glycerol, 5 β -mercaptoethanol, 3 % bromphenolblue
SyBr gold		0.01 % in 0.5 x TBE

Table 17: DNA and buffers used for EMSA

4.4.6 Crystallization of ACID variants

For crystallization trials of the ACID domain the sitting drop method was used. Commercial screens from Hampton (Hampton I, Hampton II, Natrix, Peg/Ion screen) and Natrix (classic, anion, cation, MPD) were pre-pipetted into 96 well deep well plates. These were used for setups with a Hydro plus 1 crystallization robot in Nextal conical 96 well crystallization plates. Drops of 0.5 μ l of protein (10 to 20 mg/ml) and 0.5 μ l of reservoir solution were mixed. TCEP was used as a reducing agent (0.5 μ l 1M stock solution per 50 μ l of reservoir solution).

References

- Akoulitchev, S., Chuikov, S., and Reinberg, D. (2000). TFIID is negatively regulated by CDK8-containing Mediator complexes. *Nature*, 407(6800):102–6. 0028-0836 Journal Article.
- Ansari, A. Z., Koh, S. S., Zaman, Z., Bongards, C., Lehming, N., Young, R. A., and Ptashne, M. (2002). Transcriptional activating regions target a cyclin-dependent kinase. *Proc Natl Acad Sci U S A*, 99(23):14706–9. 0027-8424 Journal Article.
- Armache, K. J., Mitterweger, S., Meinhart, A., and Cramer, P. (2005). Structures of complete RNA polymerase II and its subcomplex, Rbp4/7. *J Biol Chem*, 280(8):7131–4. 0021-9258 Journal Article.
- Asturias, F. J. (2004). RNA polymerase II structure, and organization of the preinitiation complex. *Curr Opin Struct Biol*, 14(2):121–9. 0959-440x Journal Article.
- Asturias, F. J., Jiang, Y. W., Myers, L. C., Gustafsson, C. M., and Kornberg, R. D. (1999). Conserved structures of Mediator and RNA polymerase II holoenzyme. *Science*, 283(5404):985–7. 0036-8075 Journal Article.
- Baek, H. J., Malik, S., Qin, J., and Roeder, R. G. (2002). Requirement of TRAP/Mediator for both activator-independent and activator-dependent transcription in conjunction with TFIID-associated TAFs. *Mol Cell Biol*, 22(8):2842–52. 0270-7306 Journal Article.
- Barton, G. J. (1993). Alscript: a tool to format multiple sequence alignments. *Protein Eng*, 6(1):37–40. 0269-2139 Journal Article.
- Bhoite, L. T., Yu, Y., and Stillman, D. J. (2001). The Swi5 activator recruits the Mediator complex to the HO promoter without RNA polymerase II. *Genes Dev*, 15(18):2457–69. 0890-9369 Journal Article.
- Bjorklund, S. and Gustafsson, C. M. (2004). The Mediator complex. *Adv Protein Chem*, 67:43–65. 0065-3233 Journal Article Review Review, Tutorial.
- Bjorklund, S. and Gustafsson, C. M. (2005). The yeast Mediator complex and its regulation. *Trends Biochem Sci*, 30(5):240–4. 0968-0004 Journal Article Review Review, Tutorial.
- Blazek, E. (2005). *Biochemische Darstellung funktionaler Proteome und grosser Proteinkomplexe*. PhD thesis, LMU, München.
- Bloom, H., Beier, H., and Gross, H. (1987). *Electrophoresis*, 8:93–9.

- Borggreffe, T., Davis, R., Erdjument-Bromage, H., Tempst, P., and Kornberg, R. D. (2002). A complex of the Srb8, -9, -10, and -11 transcriptional regulatory proteins from yeast. *J Biol Chem*, 277(46):44202–7. 0021-9258 Journal Article.
- Boube, M., Joulia, L., Cribbs, D. L., and Bourbon, H. M. (2002). Evidence for a Mediator of RNA polymerase II transcriptional regulation conserved from yeast to man. *Cell*, 110(2):143–51. 0092-8674 Journal Article Review Review, Tutorial.
- Bourbon, H. M., Aguilera, A., Ansari, A. Z., Asturias, F. J., Berk, A. J., Bjorklund, S., Blackwell, T. K., Borggreffe, T., Carey, M., Carlson, M., Conaway, J. W., Conaway, R. C., Emmons, S. W., Fondell, J. D., Freedman, L. P., Fukasawa, T., Gustafsson, C. M., Han, M., He, X., Herman, P. K., Hinnebusch, A. G., Holmberg, S., Holstege, F. C., Jaehning, J. A., Kim, Y. J., Kuras, L., Leutz, A., Lis, J. T., Meisterernest, M., Naar, A. M., Nasmyth, K., Parvin, J. D., Ptashne, M., Reinberg, D., Ronne, H., Sadowski, I., Sakurai, H., Sipiczki, M., Sternberg, P. W., Stillman, D. J., Strich, R., Struhl, K., Svejstrup, J. Q., Tuck, S., Winston, F., Roeder, R. G., and Kornberg, R. D. (2004). A unified nomenclature for protein subunits of Mediator complexes linking transcriptional regulators to RNA polymerase II. *Mol Cell*, 14(5):553–7. 1097-2765 Letter.
- Boyer, T. G., Martin, M. E., Lees, E., Ricciardi, R. P., and Berk, A. J. (1999). Mammalian Srb/Mediator complex is targeted by adenovirus E1A protein. *Nature*, 399(6733):276–9. 0028-0836 Journal Article.
- Brower, C. S., Sato, S., Tomomori-Sato, C., Kamura, T., Pause, A., Stearman, R., Klausner, R. D., Malik, S., Lane, W. S., Sorokina, I., Roeder, R. G., Conaway, J. W., and Conaway, R. C. (2002). Mammalian Mediator subunit Med8 is an elongin BC-interacting protein that can assemble with Cul2 and Rbx1 to reconstitute a ubiquitin ligase. *Proc Natl Acad Sci U S A*, 99(16):10353–8. 0027-8424 Journal Article.
- Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr*, 54 (Pt 5):905–21. 0907-4449 Journal Article.
- Budisa, N., Steipe, B., Demange, P., Eckerskorn, C., Kellermann, J., and Huber, R. (1995). High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in escherichia coli. *Eur J Biochem*, 230(2):788–96. 0014-2956 Journal Article.
- Bushnell, D. A., Westover, K. D., Davis, R. E., and Kornberg, R. D. (2004). Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 angstroms. *Science*, 303(5660):983–8. 1095-9203 Journal Article.

- Cantin, G. T., Stevens, J. L., and Berk, A. J. (2003). Activation domain-Mediator interactions promote transcription preinitiation complex assembly on promoter DNA. *Proc Natl Acad Sci U S A*, 100(21):12003–8. 0027-8424 Journal Article.
- Chadick, J. Z. and Asturias, F. J. (2005). Structure of eukaryotic Mediator complexes. *Trends Biochem Sci*, 30(5):264–71. 0968-0004 Journal Article Review Review, Tutorial.
- Chao, D. M., Gadbois, E. L., Murray, P. J., Anderson, S. F., Sonu, M. S., Parvin, J. D., and Young, R. A. (1996). A mammalian SRB protein associated with an RNA polymerase II holoenzyme. *Nature*, 380(6569):82–5. 0028-0836 Journal Article.
- Chaves, R. S., Herrero, P., and Moreno, F. (1999). Med8, a subunit of the Mediator CTD complex of RNA polymerase II, directly binds to regulatory elements of *suc2* and *hck2* genes. *Biochem Biophys Res Commun*, 254(2):345–50. 0006-291x Journal Article.
- Chung, W. H., Craighead, J. L., Chang, W. H., Ezeokonkwo, C., Bareket-Samish, A., Kornberg, R. D., and Asturias, F. J. (2003). RNA polymerase II/TFIIF structure and conserved organization of the initiation complex. *Mol Cell*, 12(4):1003–13. 1097-2765 Journal Article.
- Conaway, R. C., Sato, S., Tomomori-Sato, C., Yao, T., and Conaway, J. W. (2005). The mammalian Mediator complex and its role in transcriptional regulation. *Trends Biochem Sci*, 30(5):250–5. 0968-0004 Journal Article Review Review, Tutorial.
- Corden, J. L. (1990). Tails of RNA polymerase II. *Trends Biochem Sci*, 15(10):383–7. 0968-0004 Journal Article Review Review, Tutorial.
- Cosma, M. P., Panizza, S., and Nasmyth, K. (2001). Cdk1 triggers association of RNA polymerase to cell cycle promoters only after recruitment of the Mediator by SBF. *Mol Cell*, 7(6):1213–20. 1097-2765 Journal Article.
- Cramer, P. (2002). Multisubunit RNA polymerases. *Curr Opin Struct Biol*, 12(1):89–97. 0959-440x Journal Article Review Review, Tutorial.
- Cramer, P., Bushnell, D. A., and Kornberg, R. D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science*, 292(5523):1863–76. 0036-8075 Journal Article.
- Cress, W. D. and Triezenberg, S. J. (1991). Critical structural elements of the VP16 transcriptional activation domain. *Science*, 251(4989):87–90. 0036-8075 Journal Article.
- Dahmus, M. E. (1996). Reversible phosphorylation of the C-terminal domain of RNA polymerase II. *J Biol Chem*, 271(32):19009–12. 0021-9258 Journal Article Review Review, Tutorial.

- Davis, J. A., Takagi, Y., Kornberg, R. D., and Asturias, F. A. (2002). Structure of the yeast RNA polymerase II holoenzyme: Mediator conformation and polymerase interaction. *Mol Cell*, 10(2):409–15. 1097-2765 Journal Article.
- Derewenda, Z. S. (2004). Rational protein crystallization by mutational surface engineering. *Structure (Camb)*, 12(4):529–35. 0969-2126 Journal Article.
- Donaldson, L. and Capone, J. P. (1992). Purification and characterization of the carboxyl-terminal transactivation domain of Vmw65 from herpes simplex virus type 1. *J Biol Chem*, 267(3):1411–4. 0021-9258 Journal Article.
- Dotson, M. R., Yuan, C. X., Roeder, R. G., Myers, L. C., Gustafsson, C. M., Jiang, Y. W., Li, Y., Kornberg, R. D., and Asturias, F. J. (2000). Structural organization of yeast and mammalian Mediator complexes. *Proc Natl Acad Sci U S A*, 97(26):14307–10. 0027-8424 Journal Article.
- Fabrega, C., Shen, V., Shuman, S., and Lima, C. D. (2003). Structure of an mRNA capping enzyme bound to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Mol Cell*, 11(6):1549–61. 1097-2765 Journal Article.
- Fan, H. Y. and Klein, H. L. (1994). Characterization of mutations that suppress the temperature-sensitive growth of the hpr1 delta mutant of *saccharomyces cerevisiae*. *Genetics*, 137(4):945–56. 0016-6731 Journal Article.
- Flanagan, P. M., Kelleher, R. J., r., Sayre, M. H., Tschochner, H., and Kornberg, R. D. (1991). A Mediator required for activation of RNA polymerase II transcription in vitro. *Nature*, 350(6317):436–8. 0028-0836 Journal Article.
- Fondell, J. D., Ge, H., and Roeder, R. G. (1996). Ligand induction of a transcriptionally active thyroid hormone receptor coactivator complex. *Proc Natl Acad Sci U S A*, 93(16):8329–33. 0027-8424 Journal Article.
- Garcia De La Torre, J., Huertas, M. L., and Carrasco, B. (2000). Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys J*, 78(2):719–30. 0006-3495 Journal Article.
- Geiduschek, E. P. and Ouhammouch, M. (2005). Archaeal transcription and its regulators. *Mol Microbiol*, 56(6):1397–407. 0950-382x Journal Article.
- Geiger, J. H., Hahn, S., Lee, S., and Sigler, P. B. (1996). Crystal structure of the yeast TFIIA/RBP/DNA complex. *Science*, 272(5263):830–6. 0036-8075 Journal Article.
- Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Luca-Danila, A., Anderson, K., Andre, B., Arkin, A. P., Astromoff, A., El-Bakkoury, M., Bangham, R., Benito, R., Brachat, S., Campanaro, S., Curtiss, M., Davis, K., Deutschbauer, A., Entian, K. D., Flaherty, P., Foury, F., Garfinkel, D. J., Gerstein,

- M., Gotte, D., Guldener, U., Hegemann, J. H., Hempel, S., Herman, Z., Jaramillo, D. F., Kelly, D. E., Kelly, S. L., Kotter, P., LaBonte, D., Lamb, D. C., Lan, N., Liang, H., Liao, H., Liu, L., Luo, C., Lussier, M., Mao, R., Menard, P., Ooi, S. L., Revuelta, J. L., Roberts, C. J., Rose, M., Ross-Macdonald, P., Scherens, B., Schimmack, G., Shafer, B., Shoemaker, D. D., Sookhai-Mahadeo, S., Storms, R. K., Strathern, J. N., Valle, G., Voet, M., Volckaert, G., Wang, C. Y., Ward, T. R., Wilhelmy, J., Winzeler, E. A., Yang, Y., Yen, G., Youngman, E., Yu, K., Bussey, H., Boeke, J. D., Snyder, M., Philippsen, P., Davis, R. W., and Johnston, M. (2002). Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature*, 418(6896):387–91. 0028-0836 Journal Article.
- Gromoller, A. and Lehming, N. (2000a). Srb7p is a physical and physiological target of Tup1p. *Embo J*, 19(24):6845–52. 0261-4189 Journal Article.
- Gromoller, A. and Lehming, N. (2000b). Srb7p is essential for the activation of a subset of genes. *FEBS Lett*, 484(1):48–54. 0014-5793 Journal Article.
- Gu, W., Malik, S., Ito, M., Yuan, C. X., Fondell, J. D., Zhang, X., Martinez, E., Qin, J., and Roeder, R. G. (1999). A novel human Srb/Med-containing cofactor complex, SMCC, involved in transcription regulation. *Mol Cell*, 3(1):97–108. 1097-2765 Journal Article.
- Guglielmi, B., van Berkum, N. L., Klapholz, B., Bijma, T., Boube, M., Boschiero, C., Bourbon, H. M., Holstege, F. C., and Werner, M. (2004). A high resolution protein interaction map of the yeast Mediator complex. *Nucleic Acids Res*, 32(18):5379–91. 1362-4962 Journal Article.
- Guidi, B. W., Bjornsdottir, G., Hopkins, D. C., Lacomis, L., Erdjument-Bromage, H., Tempst, P., and Myers, L. C. (2004). Mutual targeting of Mediator and the TFIID kinase Kin28. *J Biol Chem*. 1083-351x Journal article.
- Hahn, S. (2004). Structure and mechanism of the RNA polymerase II transcription machinery. *Nat Struct Mol Biol*, 11(5):394–403. 1545-9993 Journal Article Review Review, Academic.
- Hall, D. B. and Struhl, K. (2002). The VP16 activation domain interacts with multiple transcriptional components as determined by protein-protein cross-linking in vivo. *J Biol Chem*, 277(48):46043–50. 0021-9258 Journal Article.
- Hallberg, M., Polozkov, G. V., Hu, G. Z., Beve, J., Gustafsson, C. M., Ronne, H., and Bjorklund, S. (2004). Site-specific Srb10-dependent phosphorylation of the yeast Mediator subunit Med2 regulates gene expression from the 2-microm plasmid. *Proc Natl Acad Sci U S A*, 101(10):3370–5. 0027-8424 Journal Article.

- Han, S. J., Lee, J. S., Kang, J. S., and Kim, Y. J. (2001). Med9/Cse2 and Gal11 modules are required for transcriptional repression of distinct group of genes. *J Biol Chem*, 276(40):37020–6. 0021-9258 Journal Article.
- Han, S. J., Lee, Y. C., Gim, B. S., Ryu, G. H., Park, S. J., Lane, W. S., and Kim, Y. J. (1999). Activator-specific requirement of yeast Mediator proteins for RNA polymerase II transcriptional activation. *Mol Cell Biol*, 19(2):979–88. 0270-7306 Journal Article.
- Hengartner, C. J., Myer, V. E., Liao, S. M., Wilson, C. J., Koh, S. S., and Young, R. A. (1998). Temporal regulation of RNA polymerase II by Srb10 and Kin28 cyclin-dependent kinases. *Mol Cell*, 2(1):43–53. 1097-2765 Journal Article.
- Hengartner, C. J., Thompson, C. M., Zhang, J., Chao, D. M., Liao, S. M., Koleske, A. J., Okamura, S., and Young, R. A. (1995). Association of an activator with an RNA polymerase II holoenzyme. *Genes Dev*, 9(8):897–910. 0890-9369 Journal Article.
- Hoepfner, S., Baumli, S., and Cramer, P. (2005). Structure of the Mediator subunit Cyclin C and its implications for CDK8 function. *J Mol Biol*, 350(5):833–42. 0022-2836 Journal Article.
- Holm, L. and Sander, C. (1995). Dali: a network tool for protein structure comparison. *Trends Biochem Sci*, 20(11):478–80. 0968-0004 Journal Article.
- Holstege, F. C., Jennings, E. G., Wyrick, J. J., Lee, T. I., Hengartner, C. J., Green, M. R., Golub, T. R., Lander, E. S., and Young, R. A. (1998). Dissecting the regulatory circuitry of a eukaryotic genome. *Cell*, 95(5):717–28. 0092-8674 Journal Article.
- Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y. (2001). A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc Natl Acad Sci U S A*, 98(8):4569–74. 0027-8424 Journal Article.
- Jiang, Y. W., Veschambre, P., Erdjument-Bromage, H., Tempst, P., Conaway, J. W., Conaway, R. C., and Kornberg, R. D. (1998). Mammalian Mediator of transcriptional regulation and its possible role as an end-point of signal transduction pathways. *Proc Natl Acad Sci U S A*, 95(15):8538–43. 0027-8424 Journal Article.
- Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A*, 47 (Pt 2):110–9. 0108-7673 Journal Article.
- Jonker, H. R., Wechselberger, R. W., Boelens, R., Folkers, G. E., and Kaptein, R. (2005). Structural properties of the promiscuous VP16 activation domain. *Biochemistry*, 44(3):827–39. 0006-2960 Journal Article.

- Junius, F. K., O'Donoghue, S. I., Nilges, M., Weiss, A. S., and King, G. F. (1996). High resolution NMR solution structure of the leucine zipper domain of the c-Jun homodimer. *J Biol Chem*, 271(23):13663–7. 0021-9258 Journal Article.
- Kadonaga, J. T. (2004). Regulation of RNA polymerase II transcription by sequence-specific DNA binding factors. *Cell*, 116(2):247–57. 0092-8674 Journal Article Review.
- Kamenski, T., Heilmeyer, S., Meinhart, A., and Cramer, P. (2004). Structure and mechanism of RNA polymerase II CTD phosphatases. *Mol Cell*, 15(3):399–407. 1097-2765 Journal Article.
- Kang, J. S., Kim, S. H., Hwang, M. S., Han, S. J., Lee, Y. C., and Kim, Y. J. (2001). The structural and functional organization of the yeast Mediator complex. *J Biol Chem*, 276(45):42003–10. 0021-9258 Journal Article.
- Kelleher, R. J., r., Flanagan, P. M., and Kornberg, R. D. (1990). A novel Mediator between activator proteins and the RNA polymerase II transcription apparatus. *Cell*, 61(7):1209–15. 0092-8674 Journal Article.
- Keller, W., Konig, P., and Richmond, T. J. (1995). Crystal structure of a bZIP/DNA complex at 2.2 Å: determinants of DNA specific recognition. *J Mol Biol*, 254(4):657–67. 0022-2836 Journal Article.
- Kettenberger, H., Armache, K. J., and Cramer, P. (2004). Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIIS. *Mol Cell*, 16(6):955–65. 1097-2765 Journal Article.
- Kim, T., Kwon, J., Kim, J., Song, Y., Kim, S., and Kim, Y. (2004). MED16 and MED23 of Mediator are coactivators of lipopolysaccharide- and heat-shock-induced transcriptional activators. *Proc. Natl. Acad. Sci. U. S. A.*, 101:12153–12158.
- Kim, Y., Geiger, J., Hahn, S., and Sigler, P. (1993). Crystal structure of a yeast TBP/TATA-box complex. *Nature*, 365:512–520.
- Kim, Y. J., Bjorklund, S., Li, Y., Sayre, M. H., and Kornberg, R. D. (1994). A multi-protein Mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell*, 77(4):599–608. 0092-8674 Journal Article.
- Kim, Y. J. and Lis, J. T. (2005). Interactions between subunits of drosophila Mediator and activator proteins. *Trends Biochem Sci*, 30(5):245–9. 0968-0004 Journal Article Review Review, Tutorial.
- Koh, S. S., Ansari, A. Z., Ptashne, M., and Young, R. A. (1998). An activator target in the RNA polymerase II holoenzyme. *Mol Cell*, 1(6):895–904. 1097-2765 Journal Article.

- Kornberg, R. D. (2005). Mediator and the mechanism of transcriptional activation. *Trends Biochem Sci*, 30(5):235–9. 0968-0004 Journal Article Review Review, Tutorial.
- Kretzschmar, M., Stelzer, G., Roeder, R. G., and Meisterernst, M. (1994). RNA polymerase II cofactor PC2 facilitates activation of transcription by Gal4-AH in vitro. *Mol Cell Biol*, 14(6):3927–37. 0270-7306 Journal Article.
- Kwon, J. Y., Kim-Ha, J., Lee, B. J., and Lee, J. (2001). The Med-7 transcriptional Mediator encoded by let-49 is required for gonad and germ cell development in *Caenorhabditis elegans*. *FEBS Lett*, 508(3):305–8. 0014-5793 Journal Article.
- La Fortelle, E. and Bricogne, G. (1997). *Meth Enzym*, B:472–494.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227(5259):680–5. 0028-0836 Journal Article.
- Lee, Y. C. and Kim, Y. J. (1998). Requirement for a functional interaction between Mediator components Med6 and Srb4 in RNA polymerase II transcription. *Mol Cell Biol*, 18(9):5364–70. 0270-7306 Journal Article.
- Levine, M. and Tjian, R. (2003). Transcription regulation and animal diversity. *Nature*, 424(6945):147–51. 1476-4687 Journal Article Review.
- Li, Y., Bjorklund, S., Jiang, Y. W., Kim, Y. J., Lane, W. S., Stillman, D. J., and Kornberg, R. D. (1995). Yeast global transcriptional regulators Sin4 and Rgr1 are components of Mediator complex/RNA polymerase II holoenzyme. *Proc Natl Acad Sci U S A*, 92(24):10864–8. 0027-8424 Journal Article.
- Li, Y. and Kornberg, R. D. (1994). Interplay of positive and negative effectors in function of the C-terminal repeat domain of RNA polymerase II. *Proc Natl Acad Sci U S A*, 91(6):2362–6. 0027-8424 Journal Article.
- Lima, C. D. (2005). Inducing interactions with the CTD. *Nat Struct Mol Biol*, 12(2):102–3. 1545-9993 Comment News.
- Lin, P. S., Dubois, M. F., and Dahmus, M. E. (2002). TFIIF-associating carboxyl-terminal domain phosphatase dephosphorylates phosphoserines 2 and 5 of RNA polymerase II. *J Biol Chem*, 277(48):45949–56. 0021-9258 Journal Article.
- Linder, T. and Gustafsson, C. M. (2004). The Soh1/MED31 protein is an ancient component of *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae* Mediator. *J Biol Chem*, 279(47):49455–9. 0021-9258 Journal Article.
- Liu, Y., Gong, W., Huang, C. C., Herr, W., and Cheng, X. (1999). Crystal structure of the conserved core of the herpes simplex virus transcriptional regulatory protein VP16. *Genes Dev*, 13(13):1692–703. 0890-9369 Journal Article.

- Liu, Y., Kung, C., Fishburn, J., Ansari, A. Z., Shokat, K. M., and Hahn, S. (2004). Two cyclin-dependent kinases promote RNA polymerase II transcription and formation of the scaffold complex. *Mol Cell Biol*, 24(4):1721–35. 0270-7306 Journal Article.
- Liu, Y., Ranish, J. A., Aebersold, R., and Hahn, S. (2001). Yeast nuclear extract contains two major forms of RNA polymerase II Mediator complexes. *J Biol Chem*, 276(10):7169–75. 0021-9258 Journal Article.
- Lorch, Y., Beve, J., Gustafsson, C. M., Myers, L. C., and Kornberg, R. D. (2000). Mediator-nucleosome interaction. *Mol Cell*, 6(1):197–201. 1097-2765 Journal Article.
- Lupas, A., Van Dyke, M., and Stock, J. (1991). Predicting coiled coils from protein sequences. *Science*, 252(5010):1162–4. 0036-8075 Journal Article.
- Lutzmann, M., Kunze, R., Buerer, A., Aebi, U., and Hurt, E. (2002). Modular self-assembly of a y-shaped multiprotein complex from seven nucleoporins. *Embo J*, 21(3):387–97. 0261-4189 Journal Article.
- Malik, S., Gu, W., Wu, W., Kin, J., and Roeder, R. G. (2000). The USA-derived transcriptional coactivator PC2 is a submodule of TRAP/SMCC and acts synergistically with other PCs. *Mol Cell*, 5:753–760.
- Malik, S. and Roeder, R. G. (2000). Transcriptional regulation through Mediator-like coactivators in yeast and metazoan cells. *Trends Biochem Sci*, 25(6):277–83. 0968-0004 Journal Article Review Review, Tutorial.
- Malik, S. and Roeder, R. G. (2005). Dynamic regulation of polymerase II by the mammalian Mediator complex. *Trends Biochem Sci*, 30(5):256–63. 0968-0004 Journal Article Review Review, Tutorial.
- Meinhart, A., Blobel, J., and Cramer, P. (2003a). An extended winged helix domain in general transcription factor TFIIE alpha. *J Biol Chem*, 278(48):48267–74. 0021-9258 Journal Article.
- Meinhart, A. and Cramer, P. (2004). Recognition of RNA polymerase II carboxy-terminal domain by 3'-RNA-processing factors. *Nature*, 430(6996):223–6. 1476-4687 Journal Article.
- Meinhart, A., Kamenski, T., Hoepfner, S., Baumli, S., and Cramer, P. (2005). A structural perspective of CTD function. *Genes Dev*, 19(12):1401–15. 0890-9369 Journal Article Review.
- Meinhart, A., Silberzahn, T., and Cramer, P. (2003b). The mRNA transcription/processing factor Ssu72 is a potential tyrosine phosphatase. *J Biol Chem*, 278(18):15917–21. 0021-9258 Journal Article.

- Meisterernst, M., Roy, A. L., Lieu, H. M., and Roeder, R. G. (1991). Activation of class II gene transcription by regulatory factors is potentiated by a novel activity. *Cell*, 66(5):981–93. 0092-8674 Journal Article.
- Mittler, G., Kremmer, E., Timmers, H. T., and Meisterernst, M. (2001). Novel critical role of a human Mediator complex for basal RNA polymerase II transcription. *EMBO Rep*, 2(9):808–13. 1469-221x Journal Article.
- Mittler, G., Stuhler, T., Santolin, L., Uhlmann, T., Kremmer, E., Lottspeich, F., Berti, L., and Meisterernst, M. (2003). A novel docking site on Mediator is critical for activation by VP16 in mammalian cells. *Embo J*, 22(24):6494–504. 0261-4189 Journal Article.
- Murakami, K. S. and Darst, S. A. (2003). Bacterial RNA polymerases: the whole story. *Curr Opin Struct Biol*, 13(1):31–9. 0959-440x Journal Article Review Review, Tutorial.
- Myers, L. C., Gustafsson, C. M., Bushnell, D. A., Lui, M., Erdjument-Bromage, H., Tempst, P., and Kornberg, R. D. (1998). The Med proteins of yeast and their function through the RNA polymerase II carboxy-terminal domain. *Genes Dev*, 12(1):45–54. 0890-9369 Journal Article.
- Myers, L. C., Gustafsson, C. M., Hayashibara, K. C., Brown, P. O., and Kornberg, R. D. (1999). Mediator protein mutations that selectively abolish activated transcription. *Proc Natl Acad Sci U S A*, 96(1):67–72. 0027-8424 Journal Article.
- Myers, L. C. and Kornberg, R. D. (2000). Mediator of transcriptional regulation. *Annu Rev Biochem*, 69:729–49. 0066-4154 Journal Article Review Review, Tutorial.
- Naar, A. M., Beaurang, P. A., Zhou, S., Abraham, S., Solomon, W., and Tjian, R. (1999). Composite co-activator ARC mediates chromatin-directed transcriptional activation. *Nature*, 398(6730):828–32. 0028-0836 Journal Article.
- Naar, A. M., Lemon, B. D., and Tjian, R. (2001). Transcriptional coactivator complexes. *Annu Rev Biochem*, 70:475–501. 0066-4154 Journal Article Review Review, Tutorial.
- Naar, A. M., Taatjes, D. J., Zhai, W., Nogales, E., and Tjian, R. (2002). Human CRSP interacts with RNA polymerase II CTD and adopts a specific CTD-bound conformation. *Genes Dev*, 16(11):1339–44. 0890-9369 Journal Article.
- Nair, D., Kim, Y., and Myers, L. C. (2005). Mediator and TFIIF govern CTD-dependent transcription in yeast extracts. *J Biol Chem*. 0021-9258 Journal article.
- Nikolov, D., Chen, H., Halay, E., Usheva, A., Hisatake, K., Lee, T., Roeder R.G., and Burley, S. (1995). Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature*, 377:119–128.

- Nonet, M. L. and Young, R. A. (1989). Intragenic and extragenic suppressors of mutations in the heptapeptide repeat domain of *Saccharomyces cerevisiae* RNA polymerase II. *Genetics*, 123(4):715–24. 0016-6731 Journal Article.
- Ottwinowski, Z. and Minor, W. (1996). Processing of x-ray diffraction data collected in oscillation mode. *Methods in Enzymology*, 276:307–326.
- Ouhammouch, M. (2004). Transcriptional regulation in archaea. *Curr Opin Genet Dev*, 14(2):133–8. 0959-437x Journal Article Review Review, Tutorial.
- Park, J. M., Gim, B. S., Kim, J. M., Yoon, J. H., Kim, H. S., Kang, J. G., and Kim, Y. J. (2001a). *Drosophila* Mediator complex is broadly utilized by diverse gene-specific transcription factors at different types of core promoters. *Mol Cell Biol*, 21(7):2312–23. 0270-7306 Journal Article.
- Park, J. M., Kim, H. S., Han, S. J., Hwang, M. S., Lee, Y. C., and Kim, Y. J. (2000). In vivo requirement of activator-specific binding targets of Mediator. *Mol Cell Biol*, 20(23):8709–19. 0270-7306 Journal Article.
- Park, J. M., Werner, J., Kim, J. M., Lis, J. T., and Kim, Y. J. (2001b). Mediator, not holoenzyme, is directly recruited to the heat shock promoter by HSF upon heat shock. *Mol Cell*, 8(1):9–19. 1097-2765 Journal Article.
- Pinhero, R., Liaw, P., Bertens, K., and Yankulov, K. (2004). Three cyclin-dependent kinases preferentially phosphorylate different parts of the C-terminal domain of the large subunit of RNA polymerase II. *Eur J Biochem*, 271(5):1004–14. 0014-2956 Journal Article.
- Rachez, C., Lemon, B. D., Suldan, Z., Bromleigh, V., Gamble, M., Naar, A. M., Erdjument-Bromage, H., Tempst, P., and Freedman, L. P. (1999). Ligand-dependent transcription activation by nuclear receptors requires the DRIP complex. *Nature*, 398(6730):824–8. 0028-0836 Journal Article.
- Ranish, J. A., Yudkovsky, N., and 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 TFIIB. *Genes Dev*, 13(1):49–63. 0890-9369 Journal Article.
- Reeves, W. M. and Hahn, S. (2003). Activator-independent functions of the yeast Mediator Sin4 complex in preinitiation complex formation and transcription reinitiation. *Mol Cell Biol*, 23(1):349–58. 0270-7306 Journal Article.
- Ren, S. and Rollins, B. J. (2004). Cyclin C/CDK3 promotes Rb-dependent G0 exit. *Cell*, 117(2):239–51. 0092-8674 Journal Article.
- Rost, B. (1996). Phd: predicting one-dimensional protein structure by profile-based neural networks. *Methods Enzymol*, 266:525–39. 0076-6879 Journal Article.

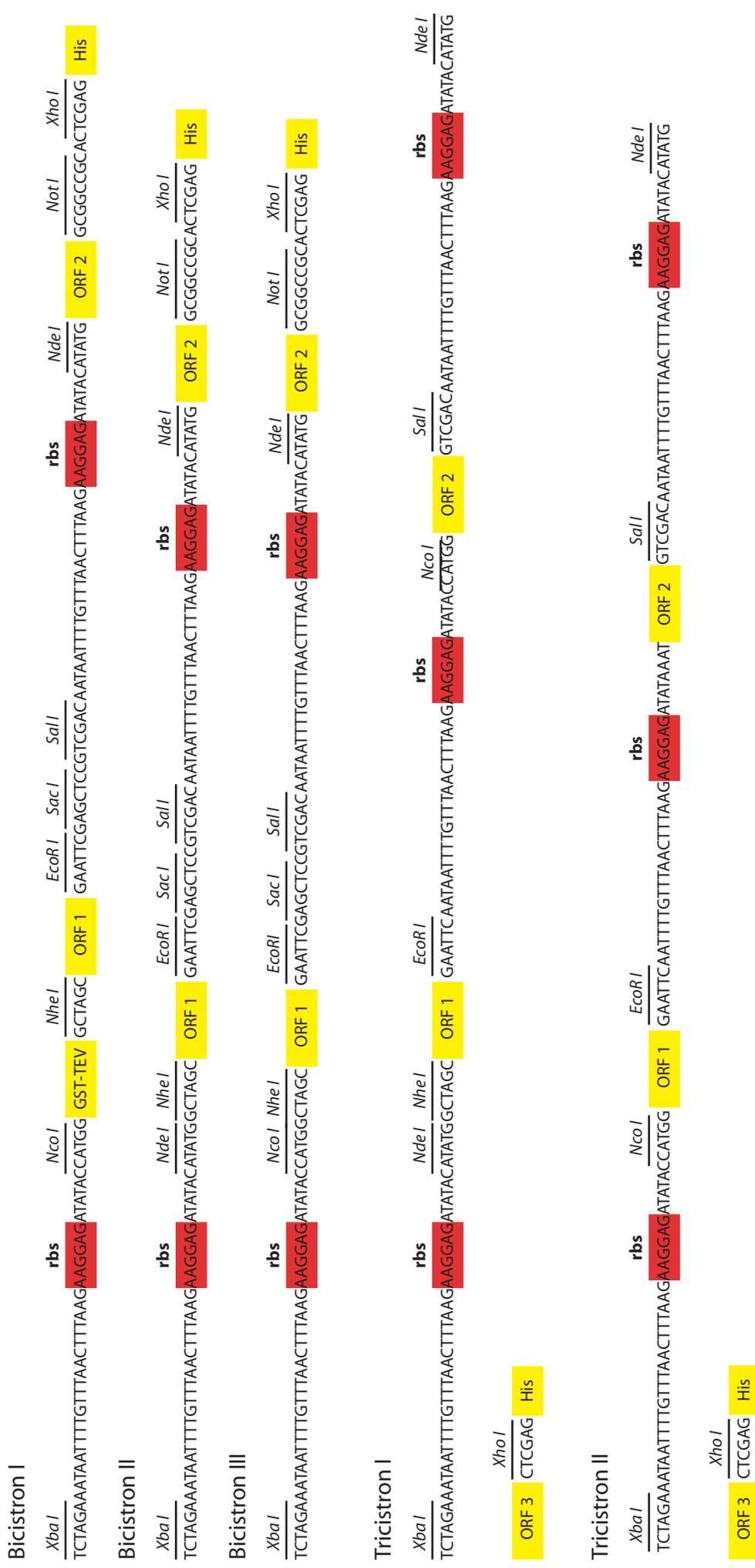
- Ryu, S. and Tjian, R. (1999). Purification of transcription cofactor complex CRSP. *Proc Natl Acad Sci U S A*, 96(13):7137–42. 0027-8424 Journal Article.
- Sadowski, I., Ma, J., Triezenberg, S., and Ptashne, M. (1988). Gal4-VP16 is an unusually potent transcriptional activator. *Nature*, 335(6190):563–4. 0028-0836 Journal Article.
- Sage, J. (2004). Cyclin C makes an entry into the cell cycle. *Dev Cell*, 6(5):607–8. 1534-5807 Journal Article Review Review, Tutorial.
- Samuelsen, C. O., Baraznenok, V., Khorosjutina, O., Spahr, H., Kieselbach, T., Holmberg, S., and Gustafsson, C. M. (2003). TRAP230/ARC240 and TRAP240/ARC250 Mediator subunits are functionally conserved through evolution. *Proc Natl Acad Sci U S A*, 100(11):6422–7. 0027-8424 Journal Article.
- Sato, S., Tomomori-Sato, C., Parmely, T. J., Florens, L., Zybaylov, B., Swanson, S. K., Banks, C. A., Jin, J., Cai, Y., Washburn, M. P., Conaway, J. W., and Conaway, R. C. (2004). A set of consensus mammalian Mediator subunits identified by multidimensional protein identification technology. *Mol Cell*, 14(5):685–91. 1097-2765 Journal Article.
- Sayre, M., Tschochner, H., and Kornberg, R. (1992). Reconstitution of transcription with five purified initiation factors and RNA polymerase II from *Saccharomyces cerevisiae*. *J. Biol. Chem.*, 267:23376–23382.
- Seipel, K., Georgiev, O., and Schaffner, W. (1994). A minimal transcription activation domain consisting of a specific array of aspartic acid and leucine residues. *Biol Chem Hoppe Seyler*, 375(7):463–70. 0177-3593 Journal Article.
- Serizawa, H., Conaway, J. W., and Conaway, R. C. (1993). Phosphorylation of C-terminal domain of RNA polymerase II is not required in basal transcription. *Nature*, 363(6427):371–4. 0028-0836 Journal Article.
- Shen, F., Triezenberg, S. J., Hensley, P., Porter, D., and Knutson, J. R. (1996). Transcriptional activation domain of the herpesvirus protein VP16 becomes conformationally constrained upon interaction with basal transcription factors. *J Biol Chem*, 271(9):4827–37. 0021-9258 Journal Article.
- Sigler, P. (1988). Transcriptional activation. Acid blobs and negative noodles. *Nature*, 333:210–2.
- Smale, S. T. and Kadonaga, J. T. (2003). The RNA polymerase II core promoter. *Annu Rev Biochem*, 72:449–79. 0066-4154 Journal Article Review.
- Spahr, H., Beve, J., Larsson, T., Bergstrom, J., Karlsson, K. A., and Gustafsson, C. M. (2000). Purification and characterization of RNA polymerase II holoenzyme from

- Schizosaccharomyces pombe. *J Biol Chem*, 275(2):1351–6. 0021-9258 Journal Article.
- Spahr, H., Khorosjutina, O., Baraznenok, V., Linder, T., Samuelsen, C. O., Hermand, D., Makela, T. P., Holmberg, S., and Gustafsson, C. M. (2003). Mediator influences Schizosaccharomyces pombe RNA polymerase II-dependent transcription in vitro. *J Biol Chem*, 278(51):51301–6. 0021-9258 Journal Article.
- Stillier, J. W. and Hall, B. D. (2002). Evolution of the RNA polymerase II C-terminal domain. *Proc Natl Acad Sci U S A*, 99(9):6091–6. 0027-8424 Journal Article.
- Struhl, K. (2005). Transcriptional activation: Mediator can act after preinitiation complex formation. *Mol Cell*, 17(6):752–4. 1097-2765 Comment Journal Article Review Review, Tutorial.
- Sullivan, S. M., Horn, P. J., Olson, V. A., Koop, A. H., Niu, W., Ebright, R. H., and Triezenberg, S. J. (1998). Mutational analysis of a transcriptional activation region of the VP16 protein of herpes simplex virus. *Nucleic Acids Res*, 26(19):4487–96. 0305-1048 Journal Article.
- Sun, X., Zhang, Y., Cho, H., Rickert, P., Lees, E., Lane, W., and Reinberg, D. (1998). NAT, a human complex containing Srb polypeptides that functions as a negative regulator of activated transcription. *Mol Cell*, 2(2):213–22. 1097-2765 Journal Article.
- Taatjes, D., Marr, T., and Tjian, R. (2004a). Regulatory diversity among metazoan co-activator complexes. *Nature Reviews*, 5:403–410.
- Taatjes, D. J., Naar, A. M., Andel, F., r., Nogales, E., and Tjian, R. (2002). Structure, function, and activator-induced conformations of the CRSP coactivator. *Science*, 295(5557):1058–62. 1095-9203 Journal Article.
- Taatjes, D. J., Schneider-Poetsch, T., and Tjian, R. (2004b). Distinct conformational states of nuclear receptor-bound CRSP-Med complexes. *Nat Struct Mol Biol*, 11(7):664–71. 1545-9993 Journal Article.
- Taatjes, D. J. and Tjian, R. (2004). Structure and function of CRSP/Med2; a promoter-selective transcriptional coactivator complex. *Mol Cell*, 14(5):675–83. 1097-2765 Journal Article.
- Takagi, Y., Chadick, J. Z., Davis, J. A., and Asturias, F. J. (2005). Preponderance of free Mediator in the yeast saccharomyces cerevisiae. *J Biol Chem*. 0021-9258 Journal article.
- Tan, S., Hunziker, Y., Sargent, D., and Richmond, T. (1996). Crystal structure of a Yeast TFIIA/TBP/DNA complex. *Nature*, 381:127–134.

- Terwilliger, T. C. (2002). Automated structure solution, density modification and model building. *Acta Crystallogr D Biol Crystallogr*, 58(Pt 11):1937–40. 0907-4449 Journal Article.
- Thompson, C. M., Koleske, A. J., Chao, D. M., and Young, R. A. (1993). A multi-subunit complex associated with the RNA polymerase II CTD and TATA-binding protein in yeast. *Cell*, 73(7):1361–75. 0092-8674 Journal Article.
- Uesugi, M., Nyanguile, O., Lu, H., Levine, A. J., and Verdine, G. L. (1997). Induced alpha helix in the VP16 activation domain upon binding to a human TAF. *Science*, 277(5330):1310–3. 0036-8075 Journal Article.
- Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., Qureshi-Emili, A., Li, Y., Godwin, B., Conover, D., Kalbfleisch, T., Vijayadamar, G., Yang, M., Johnston, M., Fields, S., and Rothberg, J. M. (2000). A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature*, 403(6770):623–7. 0028-0836 Journal Article.
- van de Peppel, J., Kettelarij, N., van Bakel, H., Kockelkorn, T. T., van Leenen, D., and Holstege, F. C. (2005). Mediator expression profiling epistasis reveals a signal transduction pathway with antagonistic submodules and highly specific downstream targets. *Mol Cell*, 19(4):511–22. 1097-2765 Journal Article.
- Verdecia, M. A., Bowman, M. E., Lu, K. P., Hunter, T., and Noel, J. P. (2000). Structural basis for phosphoserine-proline recognition by group IV WW domains. *Nat Struct Biol*, 7(8):639–43. 1072-8368 Journal Article.
- Vincent, O., Kuchin, S., Hong, S. P., Townley, R., Vyas, V. K., and Carlson, M. (2001). Interaction of the Srb10 kinase with Sip4, a transcriptional activator of gluconeogenic genes in *Saccharomyces cerevisiae*. *Mol Cell Biol*, 21(17):5790–6. 0270-7306 Journal Article.
- Walker, S., Greaves, R., and O'Hare, P. (1993). Transcriptional activation by the acidic domain of VMW65 requires the integrity of the domain and involves additional determinants distinct from those necessary for TFIIB binding. *Mol Cell Biol*, 13(9):5233–44. 0270-7306 Journal Article.
- Wang, G., Balamotis, M. A., Stevens, J. L., Yamaguchi, Y., Handa, H., and Berk, A. J. (2005). Mediator requirement for both recruitment and postrecruitment steps in transcription initiation. *Mol Cell*, 17(5):683–94. 1097-2765 Journal Article.
- Wilson, C. J., Chao, D. M., Imbalzano, A. N., Schnitzler, G. R., Kingston, R. E., and Young, R. A. (1996). RNA polymerase II holoenzyme contains SWI/SNF regulators involved in chromatin remodeling. *Cell*, 84(2):235–44. 0092-8674 Journal Article.

- Woodcock, D. M., Crowther, P. J., Doherty, J., Jefferson, S., DeCruz, E., Noyer-Weidner, M., Smith, S. S., Michael, M. Z., and Graham, M. W. (1989). Quantitative evaluation of *Escherichia coli* host strains for tolerance to cytosine methylation in plasmid and phage recombinants. *Nucleic Acids Res*, 17(9):3469–78. 0305-1048 Journal Article.
- Woychik, N. A. and Hampsey, M. (2002). The RNA polymerase II machinery: structure illuminates function. *Cell*, 108(4):453–63. 0092-8674 Journal Article Review Review, Tutorial.
- Yang, F., DeBeaumont, R., Zhou, S., and Naar, A. M. (2004). The activator-recruited cofactor/Mediator coactivator subunit ARC92 is a functionally important target of the VP16 transcriptional activator. *Proc Natl Acad Sci U S A*, 101(8):2339–44. 0027-8424 Journal Article.
- Yudkovsky, N., Ranish, J. A., and Hahn, S. (2000). A transcription reinitiation intermediate that is stabilized by activator. *Nature*, 408(6809):225–9. 0028-0836 Journal Article.
- Zhang, F., Sumibcay, L., Hinnebusch, A., and Swanson M.J. (2004). A triad of subunits from the Gall11/tail domain of Srb Mediator is an in vivo target of transcriptional activator Gcn4p. *Mol. Cell. Biol.*, 24:6871–6886.

Multiple cloning sites of multicistronic vectors



Curriculum Vitae

Sonja Baumli

Personal Data

Date of Birth: January 26, 1978
Place of Birth: Rain/LU, Switzerland
Marital status: single
Citizenship: Switzerland

Education

1984-1990 Primary School, Rain
1990-1994 Kantonsschule Beromünster
1994-1997 Kantonsschule Luzern

Academic Training

1997-2001 University of Basel, Biocenter,
Undergraduate course of Biochemistry
2000-2001 University of Basel, Biocenter,
Diploma thesis:
Structural studies on the enzyme L-Dopa
Decarboxylase
Supervisor: Prof Dr. Ueli Aebi and Dr. Peter
Burkhard
2002-2005 LMU Munich, Gene Center Graduate studies:
Structure of the MED7/MED21 heterodimer and
reconstitution of a recombinant Mediator middle
module complex
Supervisor: Prof. Dr. Patrick Cramer

23.11.05