

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

Structure of the Mediator subunit Cyclin C and subunit interaction studies within the Mediator head module



Sabine Höppner
aus Mönchengladbach, Deutschland
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 21.7.2005

Sabine Höppner

Dissertation eingereicht am

1. Gutachter: Prof. Dr. Patrick Cramer
2. Gutachter: Prof. Dr. Karl-Peter Hopfner

Mündliche Prüfung am

27.10.2005

Acknowledgements

I want to thank Patrick Cramer for everything that I learned during my PhD in his lab, especially for giving me the opportunity to work as independent as possible.

Karim, Sonja, Erika, and Hubert are the four fellows who shared this whole time with me, I'm very glad that I met them and grateful for their understanding friendship. Sonja, thanks for all your helpful contributions and sharing of ideas concerning our work.

I am especially indebted to Claudia for help and advice and for all the little things in the lab and to Toni for teaching me about crystallography. I want to thank all the other members of the Cramer lab who started and those that stayed only for a short diploma thesis or internship for contributing to the lively atmosphere, discussions, and respect, which people in this lab communicate to each other. Laurent, thanks for making it so easy to hand over my project to you by being a pleasantly uncomplicated person.

I am happy for the tea breaks that I shared with my friend Rick in the first year, I am glad that he chose to survive and recover after a hard stroke of life.

I have to thank Ralf-Peter Jansen, Katja Strässer, and Ismail Moarefi for very helpful advice as my PhD committee as well as Karl-Peter Hopfner and his whole group for the same.

I definitely need to express my thanks to the staff of Kinderkrippe Schachenmeierstrasse for ensuring that Nicolas loves to go there and for providing both sided trust.

I am grateful for having a great and supporting family that by a joined effort made it possible for me to do my work. I want to thank Nicolas for being my son and fulfilling this job in an unfussy way and Simon for being a cooperative unborn. Thank you my beloved Christoph for love and support and for believing in real sharing of duties in our partnership.

Summary

The Mediator of transcriptional regulation is the central coactivator that enables a response of RNA polymerase II to activators and repressors. It is conserved from yeast to human and consists of 25 subunits in yeast that are organized in four modules called head, middle, tail, and CDK8/Cyclin C module. Despite its central role in transcription the functional mechanism remains enigmatic. To overcome the lack of detailed structural data on the Mediator a recombinant expression system was established that allows large-scale purifications of Mediator head module subcomplexes. It has been shown that via limited proteolysis assays and multicistronic expression the problems of insolubility and low expression rates of Mediator subunits can be overcome, paving the way for structural studies on subcomplexes of the Mediator head module. First data indicated that a reconstitution of the complete head module is within close reach. Large-scale copurification data led to a detailed interaction map of subunits and subcomplexes from within the head module and towards the middle module.

The second part of this work describes the structure solution of a subunit in the CDK8/Cyclin C module – Cyclin C. Cyclin C binds the cyclin-dependent kinases CDK8 and CDK3, which regulate mRNA transcription and the cell cycle, respectively. The crystal structure of Cyclin C reveals two canonical five-helix repeats and a specific N-terminal helix. In contrast to other cyclins, the N-terminal helix is short, mobile, and in an exposed position that allows for interactions with proteins other than the CDKs. A model of the CDK8/Cyclin C pair reveals two regions in the interface with apparently distinct roles. A conserved region explains promiscuous binding of cyclin C to CDK8 and CDK3, and a non-conserved region may be responsible for discrimination of CDK8 against other CDKs involved in transcription. A conserved and Cyclin C-specific surface groove may recruit substrates near the CDK8 active site. Activation of CDKs generally involves phosphorylation of a loop at a threonine residue. In CDK8, this loop is longer and the threonine is absent suggesting an alternative mechanism of activation is discussed based on a CDK8-Cyclin C model.

Publications

The following provides a 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.

Höppner C., Carle A., Sivanesan D., Hoepfner S., Baron C. (2005)

The putative lytic transglycosylase VirB1 from *Brucella suis* interacts with the type IV secretion system core components VirB8, VirB9 and VirB11.

Submitted.

Table of Contents

INTRODUCTION	1
1. Transcription and the Mediator	1
1.1. Gene transcription is accomplished by RNA polymerases	1
1.2. Initiation – starting the transcription cycle	2
1.3. During the transcription cycle RNA Pol II undergoes regulatory phosphorylation and dephosphorylation	3
1.4. Transcription needs regulation	3
1.5. Discovery of a Mediator of transcriptional regulation in yeast	4
1.6. Mediator functions in transcription initiation and reinitiation	5
1.7. Architecture of the Mediator	7
1.8. A unified nomenclature for Mediator proteins	9
2. The Mediator CDK8/Cyclin C Module	11
2.1. CDK-cyclin pairs	11
2.2. Three CDKs differentially phosphorylate the CTD and regulate transcription	12
2.2.1. CDK7/Cyclin H	12
2.2.2. CDK8/Cyclin C	12
2.2.3. CDK9/Cyclin T	14
2.2.4. Substrate specificity and regulation of CTD kinases	15
3. Aims of this work	18
MATERIALS AND METHODS	20
1. Bacterial strains and insect cells	20
2. Plasmids and genomic DNA	20
3. Media	23
4. Buffers and solutions	24
5. Molecular cloning techniques	26
5.1. Oligonucleotides	26
5.2. PCR	26
5.3. DNA isolation	27
5.4. Restriction cleavage and dephosphorylation	27
5.5. Ligation	27
5.6. Mutagenesis	27
5.7. Transformation of plasmid DNA	27
5.8. Preparation of competent cells	28
6. Biochemical methods	28
6.1. Protein expression and purifications	28
6.1.1. Protein expression in <i>E. coli</i> cells	28
6.1.2. Cell lysis and chromatography	28
6.1.2.1. Cell lysis and affinity chromatography	28
6.1.2.2. Ion exchange chromatography	29
6.1.2.3. Gel filtration	29
6.1.3. Enrichment of proteins	29
6.1.4. Individual purification protocols	29
6.1.4.1. Purification of GST-MED17 (GST-Srb4)	30
6.1.4.2. Purification of GST-MED17-F2 (GST-Srb4-F2)	31
6.1.4.3. Purification of MED17-F1_His ₆ (Srb4-F1_His ₆)	31
6.1.4.4. Purification of MED17-F2_His ₆ -bicistrons	32

6.1.4.5.	Purification of the GST-MED17-F2/MED6 Δ C_His ₆	32
6.1.4.6.	Purification of MED20/MED18_His ₆ (Srb2/Srb5_His ₆) and MED8/MED20/MED18_His ₆	33
6.1.4.7.	Purification of the GST-MED17-F2/MED6 Δ C MED8/MED20/MED18-His ₆ coexpression	34
6.1.4.8.	Purification of CDK8_His ₆	34
6.1.4.9.	Purification of Cyclin C	35
6.2.	Limited Proteolysis	35
6.3.	Determination of protein-protein interactions	36
6.3.1.	GST Sepharose® pull-down assay	36
6.3.2.	Microcalorimetry	37
7.	SF9 insect cell culture and recombinant baculovirus expression techniques	37
7.1.	Growth conditions for SF9 cells	37
7.2.	Freezing and thawing of insect cells	37
7.3.	Transposon mutagenesis and blue/white selection	38
7.4.	Isolation of Bacmid DNA	38
7.5.	Transfection of SF9 cells with Bacmid DNA	38
7.6.	Harvesting of initial virus stocks and virus reamplification	39
7.7.	Expression of CDK8	39
8.	Electrophoretic methods	39
8.1.	Electrophoretic separation of DNA	39
8.2.	Protein separation by SDS-PAGE	40
9.	Immunological methods	40
9.1.	Protein transfer and Western blot	40
9.2.	Passive adsorption method for protein transfer	40
10.	Crystallization	41
11.	X-Ray analysis	41
11.1.	Data collection	41
12.	Data processing and structure solution	42
12.1.	Indexing, integration and scaling	42
12.2.	Phasing	42
12.3.	Refinement	42
RESULTS		44
1.	Mapping of subunit domains and subunit-subunit interactions within the Mediator head module	44
1.1.	MED17 (Srb4), the integral subunit of the Mediator head module	44
1.2.	Two fragments of MED17 (Srb4) display better expression and solubility than full-length protein	47
1.3.	Bicistronic expressions of MED17 (Srb4) with the other Mediator head module subunits reveals binding to MED6 and MED8	47
1.3.1.	Screening with bicistronic vectors reveals binding of MED17 to MED8	48
1.3.2.	MED6 bridges the two Mediator core modules	49
1.4.	A trimeric subcomplex MED20/MED18/MED8 in the Mediator head module	50
1.4.1.	Preparation of soluble MED20/MED18 (Srb2/Srb5) - heterodimer	50
1.4.2.	Crystallization of MED20/MED18 (Srb2/Srb5) heterodimer	51
1.4.3.	MED8 binds to the MED20/MED18 (Srb2/Srb5) heterodimer	52
1.4.4.	Limited proteolysis reveals potential variants of MED18 (Srb5) and confirms the high sensitivity and instability of the MED8 protein	53
1.5.	Design of a new tricistronic vector	54
1.6.	MED18 binds to MED8	55

1.7.	Coexpression of the dimeric MED6/MED17-F2 (Srb4-F2) and the trimeric MED20/MED18/MED8 (Srb2/Srb5/MED8) complex results in a pentameric head module subcomplex	56
1.8.	Overview on discovered contacts in the Mediator head module	57
1.9.	Binding assays point towards interactions between Mediator subcomplexes and RNA Pol II	57
2.	Structural studies of the Mediator CDK8/Cyclin C module	59
2.1.	Cyclin C and CDK8 from <i>S. pombe</i> are obtained by recombinant expression techniques	59
2.1.1.	BLAST search discovery and cloning of Cyclin C and CDK8	59
2.1.2.	CDK8 expression with the baculovirus <i>Bac-to-Bac</i> system	60
2.1.3.	Cyclin C can be recombinantly expressed in <i>E. coli</i> and purified to crystallizable amounts	62
2.1.3.1.	Limited proteolysis - the protease can make the difference	63
2.1.3.2.	Two leucines in Cyclin C were mutated to methionines to allow MAD phasing	64
2.2.	Solution of the Mediator Cyclin C structure	65
2.2.1.	Data collection	65
2.2.2.	Phasing and refinement	66
2.3.	Analysis of the Cyclin C structure	68
2.3.1.	The structure of Cyclin C - canonical cyclin repeats	68
2.3.2.	A mobile N-terminal helix	70
2.3.3.	Structure-based alignments.	70
2.3.4.	Modeling of the CDK8-Cyclin C pair.	72
2.3.5.	Aligning CDK classes within the protein family allows determination of CDK8 characteristics and their mapping on the CDK8 model	72
2.3.6.	Specificity of CDK-cyclin interaction.	74
2.3.7.	Conserved surfaces.	77
2.3.8.	A conserved groove unique to Cyclin C.	77
2.3.9.	Possible mechanisms of CDK8 activation.	78
	DISCUSSION AND FUTURE DIRECTIONS	80
1.	The Mediator head module: dissecting a protein-protein network	80
1.1.	MED17 – domains and interactions	80
1.2.	The MED20/MED18/MED8 proteins form a stable subcomplex within the Mediator head module	82
1.3.	Towards reconstitution of the Mediator head module and the Mediator core	83
1.4.	Structure-function studies on Mediator-RNA Pol II complexes	85
2.	The Mediator CDK8/Cyclin C module: implications of the Cyclin C structure for function	86
2.1.	The Cyclin C structure	86
2.2.	Cyclin C differs from known cyclin structures	88
2.3.	A structural model of CDK8 identifies targets for functional analysis	89
2.4.	Analysis of CDK-Cyclin interface suggests the structural basis for cyclin promiscuity	90
2.5.	A highly specific groove of Cyclin C	91
2.6.	The CDK8/Cyclin C module and its role in Mediator function	92
	REFERENCES	94
	SUPPLEMENTARY MATERIAL	111
	CURRICULUM VITAE	113

Introduction

1. Transcription and the Mediator

1.1. Gene transcription is accomplished by RNA polymerases

RNA synthesis, or transcription, is the process of transcribing DNA nucleotide sequence information into RNA sequence information. RNA synthesis is catalyzed by a large enzyme, the RNA polymerase. The process of gene transcription functions in a very similar way in all life forms, thus was conserved over more than a billion years of evolution. However, even though RNA synthesis itself, which is always catalyzed by an RNA polymerase, might be quite similar – the central enzymes of all known organisms share homologies (Ebright, 2000) – substantial differences are made in the kingdoms of life on how to organize and regulate this process. In prokaryotes one single RNA polymerase transcribes all RNA, whereas in eukaryotes (fungi and metazoans) this task is carried out by three polymerase paralogs. RNA polymerase I transcribes rRNA and RNA polymerase III tRNAs. All protein-coding genes are transcribed by RNA polymerase II (RNA Pol II).

RNA synthesis takes place in three stages: initiation, elongation and termination. RNA synthesis is initiated at specific DNA sequences, so called promoters. The prokaryotic *E. coli* RNA polymerase holoenzyme is a big 400 kD core complex comprising 5 subunits ($2 \times \alpha$, β , β' , σ), the initiation specific σ -subunit can dissociate from this holoenzyme, leaving a four subunit core enzyme (Young *et al.*, 2002). The core enzyme is responsible for template-directed RNA synthesis and also interacts with regulatory proteins, which modulate transcription levels. The σ -subunit is responsible for the site-specific recognition of the promoter element as well as melting of the promoter DNA. It is thought to dissociate and stay at the site of transcription, facilitating transcription initiation by the next core enzyme. However, new results implicate a role for the σ -factor in elongation as well (Brodolin *et al.*, 2004; Mukhopadhyay *et al.*, 2001). An additional level of gene regulation is introduced by a multiplicity of σ -factors that an *E. coli* cell possesses, each responsible for a different set of promoters and genes.

Analogous to the prokaryotic system, transcription in eukaryotic cells is initiated at promoters. Promoters for RNA Pol II typically contain a core sequence

element, to which transcription factors bind and form a nucleation site for transcription complex formation. The best-characterized, as well as most common promoters in eukaryotes carry a TATA box as the core sequence element. The RNA Pol II core enzyme consists of 12 subunits (513 kD). The structure of the RNA Pol II core enzyme has been solved (Armache *et al.*, 2005; Cramer *et al.*, 2001) and proved that RNA polymerases share a conserved core and common transcription mechanism. The eukaryotic initiation factors, the general transcription factors (GTFs), are more distantly related to the bacterial σ -factor but function in a similar manner in promoter recognition, promoter melting, abortive initiation and promoter escape (Kornberg, 2005).

1.2. Initiation – starting the transcription cycle

The beginning of a round of transcription is marked by recruitment of the transcription machinery to a promoter. The core promoter serves to position RNA Pol II in the preinitiation complex (PIC). The bacterial equivalent to the PIC is called closed complex.

Whereas in prokaryotes the σ -factor is the only additional polypeptide required for initiation, a number of additional factors are needed along with the core enzyme in order to recognize the core element and initiate transcription in eukaryotes (Hahn, 2004). *In vitro* reconstitution of eukaryotic transcription by supplementation with crude cell extracts, led to the identification of these factors (Weil *et al.*, 1979). Since factors from the crude extract had to be essential for transcription, fractionation and purification of the extract finally led to the identification of 5 essential GTFs: TFIIB, TFIID, TFIIE, TFIIIF, and TFIIH. A simplified model on how they enable transcription in eukaryotes was established (Buratowski, 1994). The first step is promoter recognition by TFIID — a multisubunit complex containing the TATA-binding protein (TBP) and at least 14 tightly associated factors (TAFs) (Albright and Tjian, 2000; Orphanides *et al.*, 1996). Binding of TFIID causes bending of the DNA and serves as an initial nucleation platform for further factors. TFIIB then binds to the platform and stabilizes the complex by contacting TFIID and sequences flanking the TATA box on both sides. TFIIB, in turn, recruits the RNA Pol II–TFIIIF complex. However, transcription cannot occur until TFIIE and TFIIH are incorporated, the PIC is formed. ATP-dependent promoter melting leads to formation of the unstable open complex where the DNA double helix is melted (“transcription bubble”) to facilitate the initiation

of transcription. This step is catalyzed by the helicase activity of TFIIH. TFIIH as well as TFIIIE are also required for promoter escape of RNA Pol II (Hahn, 2004; Koleske and Young, 1995; Kornberg, 2005; Thompson *et al.*, 1993).

1.3. During the transcription cycle RNA Pol II undergoes regulatory phosphorylation and dephosphorylation

The largest subunit of RNA Pol II comprises a carboxy-terminal repeat domain (CTD) that consists in heptapeptide repeats of the consensus sequence $Y_1S_2P_3T_4S_5P_6S_7$. The CTD is unique to RNA Pol II. The number of repeats is species dependent, and is 26 in yeast and 52 in human. At least eight repeats are required for yeast viability (Nonet *et al.*, 1987; West and Corden, 1995). The CTD, in particular S_2 and S_5 , is the target for regulatory phosphorylation and dephosphorylation during the transcription cycle (Bentley, 2002). RNA Pol II is recruited to the promoter in a hypophosphorylated state and becomes phosphorylated at S_5 during transcription initiation. Phosphorylation destabilizes the PIC leading to open complex formation and promoter escape. During elongation Pol II becomes dephosphorylated on S_5 and phosphorylated on S_2 . Dephosphorylation on S_2 enables entry into a new round of transcription. The CTD acts as a platform for assembly of factors that regulate transcription initiation, elongation, termination and mRNA processing (Hahn, 2004). The central role of the CTD in mRNA synthesis is reflected by the fact that many factors of the RNA Pol II transcription machinery bind the CTD, which is unique to RNA Pol II.

1.4. Transcription needs regulation

In order to react to stimuli from the environment, a single cell needs to regulate the production of proteins from a given set of genes. Cells in a multi-cellular organism contain the same set of genes, but the subset of genes that are expressed differentiates from one cell to another. Therefore gene expression must be regulated in order to produce the right set of proteins in correct amounts at the right instant. Regulation of gene expression is far more complex in eukaryotes than in prokaryotes and takes place at many different levels, including transcription, RNA processing, mRNA export, mRNA and protein stability and translation but also just recently

discovered processes like RNAi and many more. The focal point in the regulation of most genes is however the initiation of transcription at promoters.

RNA Pol II and GTFs alone can constitute *in vitro* basal transcription (Conaway and Conaway, 1993; Roeder, 1996; Roeder, 1998). Transcriptional regulation is achieved by elements both proximal and distal to the promoter that can bind to regulatory factors and function in activation or repression of the corresponding gene. A plethora of signals coming from the environment and manifesting themselves in a variety of different activation and repression factors somehow needs to be integrated and reduced to a simple switch for telling the transcription machinery to start or to stay at the promoter. Thus a physical interaction is required between a variety of signaling factors and a transcription machinery that is highly conserved throughout the eukaryotic kingdom. In prokaryotic systems transcriptional repressor and activator proteins bind to DNA sequences that are adjacent to promoters and exert effects directly on RNA polymerase. While repressors block binding of the polymerase to the promoter, activators increase its affinity or stimulate open complex formation. In contrast to the bacterial system regulatory proteins do not target the RNA polymerase directly (Kornberg, 2005). The huge assembly of RNA Pol II and general transcription factors is unable to support activated transcription by DNA-binding transcriptional activators *in vitro* (Myers and Kornberg, 2000). This has led to the suggestion of intermediary factors, which function as co-activators and are required for transmitting the variety of regulatory signals produced in the cell from activators and repressors to the Pol II initiation apparatus. The cofactors, co-activators as well as co-repressors, are thought to assume this adapting role. They are distinct from GTFs since basal transcription *in vitro* can function without their presence. In contrast to activators and repressors they do not bind DNA with sequence specificity themselves (Roeder, 1998).

1.5. Discovery of a Mediator of transcriptional regulation in yeast

The Existence of an intermediary complex was first suggested when an inhibitory effect called squelching was discovered in *in vitro* transcription assays. One gene activator protein interfered with the effects of another in eukaryotic cells (Gill and Ptashne, 1988). The effect remained even in the presence of large amounts of GTFs. The target of activators could thus not be identical with the known basal

transcription machinery. Interference could be reversed by a partially purified yeast component, but not by RNA polymerase II or various polymerase II transcription factors that was from crude extracts. It was suggested to be a novel factor, which mediates the transcriptional activation process (Flanagan *et al.*, 1991; Kelleher *et al.*, 1990).

Independently a genetic study led to the identification of mutations that suppress the temperature sensitive phenotype of CTD truncations. They were named *srb* for suppressor of RNAP B (Nonet and Young, 1989). At a later stage a complex was discovered containing RNA Pol II, general transcription factors and some of the *Srb* proteins. This complex was then termed the RNA Pol II holoenzyme, which could support activated transcription in an *in vitro* system when supplemented with the missing GTFs (Koleske and Young, 1994; Thompson *et al.*, 1993). The different activities and complexes that had been purified could be linked to one another when a 20 subunit Mediator to homogeneity. The connection with the *Srb*- complex was made by the finding that MED 20 (*Srb*2), (MED17) *Srb*4, (MED18) *Srb*5, and (MED22) *Srb*6 were subunits of the pure Mediator (Kim *et al.*, 1994).

Upon further purification the complex was shown to contain additional *Srb* proteins, *Srb*8-11, now known as the CDK8/Cyclin C module (Liao *et al.*, 1995).

Important functions have been ascribed to the Mediator: support of transcriptional activation, stimulation of basal transcription and stimulation of TFIIH dependent phosphorylation of the CTD (Kim *et al.*, 1994; Myers *et al.*, 1998). Later Mediator was also shown to possess a histone acetyltransferase (HAT) activity (Lorch *et al.*, 2000). In fact, Mediator is needed for almost all RNA Pol II dependent transcription *in vivo* (Holstege *et al.*, 1998). It is a still ongoing discussion how the preinitiation complex is assembled, which factors are present, and at which state RNA Pol II is recruited.

1.6. Mediator functions in transcription initiation and reinitiation

It is not clear whether recruitment and assembly of the transcription machinery at the promoter follows a sequential mode (Koleske and Young, 1995). Another model is the recruitment of two complexes, a TBP containing complex and a holoenzyme containing the remaining GTFs (Myer and Young, 1998). A spectrum of possibilities might be realized at different promoters. It is known that transcription initiation involves the interaction of the unphosphorylated CTD with the Mediator

complex (Bjorklund and Gustafsson, 2004; Bjorklund and Gustafsson, 2005; Malik and Roeder, 2000; Myers and Kornberg, 2000). The CTD is required for the formation of stable RNA Pol II-Mediator complexes in yeast (Asturias *et al.*, 1999; Myers *et al.*, 1998). Antibodies against the unphosphorylated CTD compete with this interaction and therefore displace Mediator from Pol II (Kim *et al.*, 1994; Svejstrup *et al.*, 1997). Consistent with its important central role in transcription initiation, binding of Mediator to the CTD could also be shown for metazoan organisms (Jiang *et al.*, 1998; Naar *et al.*, 2002; Park *et al.*, 2001). The CTD-Mediator interaction is required for Mediator function, since yeast Mediator cannot stimulate transcription by a CTD-less Pol II (Myers *et al.*, 1998).

The yeast Mediator binds cooperatively with Pol II and a subset of general transcription factors at an intermediary state of PIC formation (Ranish *et al.*, 1999). In biochemical fractionation approximately 40% of Mediator forms a stable complex with Pol II. On one hand this is consistent with studies showing that Mediator can be recruited to promoters independent of the rest of the transcription machinery (Cosma *et al.*, 1999; Rani *et al.*, 2004). On the other hand this might mean that RNA Pol II is recruited to the promoter as the holoenzyme, with Mediator.

After initiation of Pol II transcription *in vitro*, many of the general transcription factors remain behind at the promoter in the so-called scaffold complex (Yudkovsky *et al.*, 2000). The phosphorylated CTD does not associate with Mediator (Svejstrup *et al.*, 1997), and CTD phosphorylation during transcription initiation apparently also breaks the Pol II-Mediator interaction, resulting in an elongating polymerase and a scaffold complex where Mediator remains at the promoter (Liu *et al.*, 2004). This complex presumably marks genes that have been transcribed and enables the typically slow step of recruitment to be bypassed in subsequent rounds of transcription. Certain transcription activation domains can stabilize this scaffold complex *in vitro*. The scaffold complex can then rapidly recruit the remaining general transcription factors to promote transcription initiation.

Electron microscopic images of the yeast Pol II-Mediator complex suggest that Mediator binds tightly to RNA Pol II at multiple interaction sites (Asturias *et al.*, 1999), presumably mainly the Rpb3 and Rpb11 subunits, on the side opposite to the active site cleft (Davis *et al.*, 2002). Figure 1 depicts the relative sizes of Mediator and RNA Pol II and their suggested relative orientation. In addition, Mediator binds to several general transcription factors involved in initiation (Kang *et al.*, 2001; Park *et al.*,

2001). Nevertheless, it remains enigmatic how the large 1MDa Mediator fits into the context of the rest of the transcription machinery and by which mechanism it transmits signals from regulatory factors to Pol II.

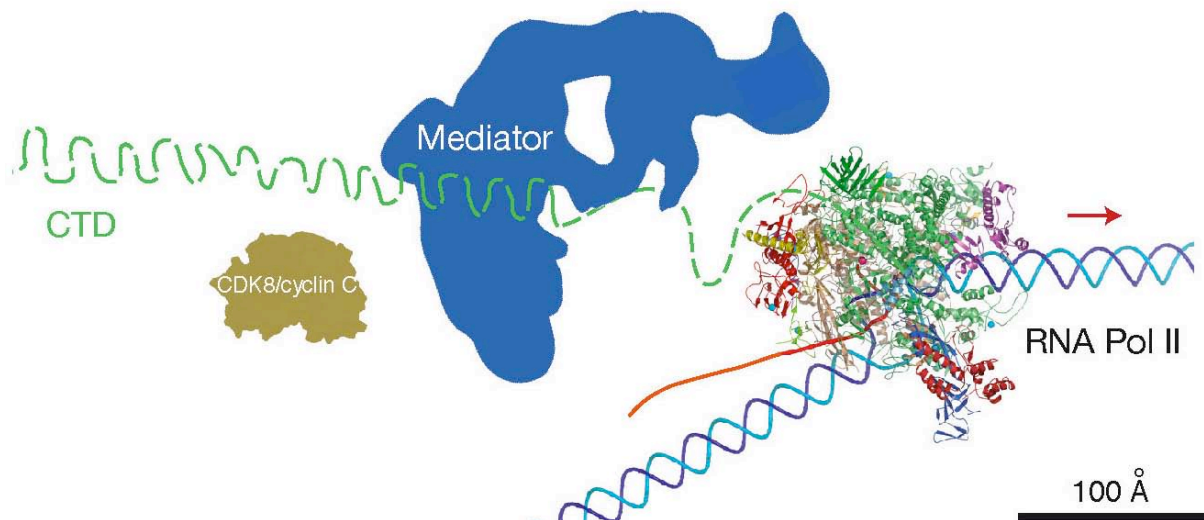


Fig.1 – RNA polymerase II and the relative size of Mediator

The complete 12-subunit yeast Pol II elongation complex structure (Armache *et al.*, 2005; Kettenberger *et al.*, 2004) is shown as a ribbon diagram with the subunits in different colors. DNA is in blue and RNA is in red. DNA and RNA outside the polymerase have been extrapolated or modeled. The relative sizes of the yeast Pol II largest subunit linker region and the CTD are indicated in green. The CTD is depicted in the putative state of a loose spiral. Projections of Mediator consisting of head, middle, and tail module and of a CDK/cyclin pair are drawn to scale.

1.7. Architecture of the Mediator

The Mediator subunit architecture was inferred from biochemical, genetic and electron microscopic studies. Mediator was originally divided in two structural submodules, MED17 (Srb4) and MED14 (Rgr1), after biochemical studies with urea dissociation (Lee and Kim, 1998). Since a knockout of *sin4* results in the loss of three subunits including MED15 (Gal11) (Myers *et al.*, 1998), the MED14 (Rgr1) module was subdivided into the MED9/10 (Nut2) and the MED15 (Gal11) module (Kang *et al.*, 2001). The MED17, MED9/19 and MED15 modules were tentatively correlated with three density lobes in electron microscopic images, termed the head, middle, and tail module, respectively (Dotson *et al.*, 2000).

The biochemically defined MED15 (Gal11) module could be assigned to the tail using again the Sin4 knockout (Dotson *et al.*, 2000). MED15 (Gal11) and MED10 (Nut2) both interact with MED14 (Rgr1), it was thus clear that the MED10 (Nut2)

protein must belong to the middle domain (Li *et al.*, 1995). As an indirect conclusion, the MED17 (Srb4) module was assigned to be the head module.

EM pictures suggest that the head module is the major RNA Pol II-interacting subcomplex while the middle module seems to make fewer contacts. The head and middle modules form a core Mediator that can be isolated from yeast (Liu *et al.*, 2001; Myers *et al.*, 1998). Core Mediator enables 4-fold activation of transcription in nuclear extracts, compared to 18-fold activation for the complete Mediator. The CTD may bind between the head and middle modules, since the Srb proteins are distributed among head and middle module and since recombinant head and middle modules independently bind to the CTD (Kang *et al.*, 2001).

The EM pictures do apparently not contain the CDK8/Cyclin C module. The additional Srb proteins 8-11 (Table 1) (Liao *et al.*, 1995; Nonet and Young, 1989) can be found in some holoenzyme preparations (Myer and Young, 1998). However, its presence depends on the preparation protocol, it has been suggested to be associated to present in the holoenzyme of cells growing exponentially and absent from cells in the stationary phase (Hengartner *et al.*, 1998).

Various lines of evidence suggest that the tail module – MED15 (Gal11), MED2 and MED3 (Pgd1) and Sin4 (Myers *et al.*, 1998) (Table1) – is the main target for transcriptional activators. The holoenzyme isolated from gal11 null mutant lacks the entire tail MED15 (GAL11) module and is functionally defective for activated, but not basal transcription (Han *et al.*, 1999; Park *et al.*, 2000) Binding assays *in vitro* show binding of activators to Mediator through the MED15 (GAL11) module (Lee *et al.*, 1999b).

Half of the subunits in the MED17 (Srb4) subcomplex, that is thought to form the head module of the Mediator, consist in Srb-proteins, namely MED20 (Srb2), MED17 (Srb4), MED18 (Srb5), MED22 (Srb6). The other four subunits are named Med6, Med8, MED19 (Rox3) and Med11 (Table 1). The 8 subunits interact with one another as shown by a variety of *in vitro* experiments (Kang *et al.*, 2001; Lee and Kim, 1998). In particular it was shown that it is possible to purify the MED17 (Srb4) complex after recombinant expression in a baculoviral system (Koh *et al.*, 1998).

The most conserved part of Mediator is the middle module (Baumli *et al.*, 2005), which comprises subunits MED7, MED21, MED10, MED1, MED4, MED9, and possibly MED31 (Table1). With the exception of MED31 (Soh1), that was shown to

be part of the Mediator only recently (Guglielmi *et al.*, 2004; Linder and Gustafsson, 2004), MED7 and MED21 show the highest degree of sequence homology of all core Mediator subunits (40% and 45% between yeast and human, respectively). This high conservation is reflected in their essential function in yeast (Myers *et al.*, 1998) and in a requirement for MED21 in mouse development (Tudor *et al.*, 1999).

Electron microscopy also showed that Mediator undergoes strong structural changes upon interaction with Pol II (Asturias *et al.*, 1999; Davis *et al.*, 2002; Naar *et al.*, 2002) and with transcription activators (Taatjes *et al.*, 2002) that are obvious even at 35 Å resolution (Davis *et al.*, 2002). The causes and consequences of these changes are however poorly understood, and the molecular mechanism of Mediator remains enigmatic. Recently the structure of the MED7/Srb7 heterodimer in the middle module was solved (Baumli *et al.*, 2005). The important role that the huge conformational changes might have for function of Mediator is reflected in highly conserved elements that confer flexibility. Such hinges could provide a mechanistic principle, which could underlie Mediator function.

1.8. A unified nomenclature for Mediator proteins

Counterparts of the *S. cerevisiae* Mediator have been identified in many higher organisms. Biochemical and bioinformatical methods finally led to understanding that Mediator is conserved from yeast to human (Boube *et al.*, 2002; Bourbon *et al.*, 2004). Several Mediator complexes have been identified from mammalian cells: thyroid hormone receptor-associated proteins (TRAP) (Fondell *et al.*, 1996; Fondell *et al.*, 1999), suppressor of RNA polymerase B (SRB)-mediator containing cofactor (SmCC) (Gu *et al.*, 1999), activator-recruited cofactor (ARC) (Naar *et al.*, 1999), vitamin D receptor-interacting proteins (DRIP) (Rachez *et al.*, 1999), negative regulator of activation transcription (NAT) (Sun *et al.*, 1998), positive cofactor 2 (PC2) (Kretzschmar *et al.*, 1994; Meisterernst *et al.*, 1991), and cofactor required for Sp1 activation (CRSP) (Ryu and Tjian, 1999; Ryu *et al.*, 1999). Tandem mass spectrometry was used to systematically identify the proteins present in the highly purified complexes and to differentiate between 30 MED subunits (MEDs) (Sato *et al.*, 2003; Tomomori-Sato *et al.*, 2004). While initially only 8 subunits of Mediator were thought to be conserved over evolution (Malik and Roeder, 2000; Rachez and Freedman, 2001) it became clear after extensive cross species comparisons that counterparts for nearly every yeast Mediator subunit could be

found in metazoans (Borggreffe *et al.*, 2002; Boube *et al.*, 2002; Gustafsson and Samuelsson, 2001; Samuelsen *et al.*, 2003). Consequently in 2004 it was proposed for reasons of simplicity and easier understanding to introduce a unified nomenclature for Mediator subunits in all species, acknowledging conservation from yeast to human. All known subunits are named from MED1 to MED31, except CDK8 and CycC (Cyclin C) (table 1). The original names however carry implications for the function of these proteins.

Table 1- Unified nomenclature of the Mediator subunits

New name in all species	<i>S. cerevisiae</i> -previous name	New name in all species	<i>S. cerevisiae</i> -previous name
Head		Tail	
MED6	Med6	MED2	Med2
MED8	Med8	MED3	Pgd1/Hrs1/Med3
MED11	Med11	MED5	Nut1
MED17	Srb4	MED15	Gal11
MED18	Srb5	MED 16	Sin4
MED 19	Rox3		
MED20	Srb2	CDK8/Cyclin C	
MED22	Srb6	CDK8	Srb10/Ssn3/Ume5
		CycC	Srb11/Ssn8/Ume3
Middle		MED12	Srb8
MED1	Med1	MED13	Ssn2/Srb9
MED4	Med4		
MED7	Med7		
MED9	Cse2/Med9		
MED10	Nut2/Med10		
MED14	Rgr1		
MED21	Srb7		
MED31	Soh1		

2. The Mediator CDK8/Cyclin C Module

2.1. CDK-cyclin pairs

Cyclin-dependent kinases (CDKs) associate with specific cyclins and play central roles in cell cycle regulation and transcription (Bregman *et al.*, 2000; Dynlacht, 1997; Murray, 2004). Phosphorylation is the major posttranslational modification of the CTD *in vivo*. The correct phosphorylation status of the CTD of RNA Pol II is required for initiation, repression and successful elongation of transcription (Bensaude *et al.*, 1999; Cadena and Dahmus, 1987; Kang and Dahmus, 1993; Lu *et al.*, 1992) but also plays an important role in mRNA capping, mRNA polyadenylation, and mRNA export. It is likely that the CTD recruits components of the splicing machinery to the RNA polymerase II elongation complex (Shilatifard *et al.*, 2003).

The specific phosphorylation status of the CTD is achieved by a number of kinases and phosphatases and big effort has been put in investigating the nature of the CTD kinases. While a large number of kinases was identified to phosphorylate the CTD *in vitro*, one group of kinases has been well established so far to be of crucial functional relevance in controlling/running the transcription cycle *in vivo*: the highly conserved paralogs of CDK7, CDK8 and CDK9, which form part of TFIIH, of Mediator, and of the Positive Transcription Elongation Factor b (P-TEFb) respectively, are implicated in key roles of the transcription cycle.

CDKs were originally found to be a family of kinases whose activity and thereby control over the cell cycle is driven by the fluctuating abundance of their regulatory cyclin subunit (Lees, 1995; Morgan, 1995). They all share a high degree of conservation that is reflected in a very similar fold (Lolli *et al.*, 2004; Russo *et al.*, 1996; Russo *et al.*, 1998; Tarricone *et al.*, 2001). However, conservation levels differ more strongly in the cyclin subunits. The CTD kinases CDK7, CDK8 and CDK9 associate with C- type cyclins, a subfamily that was first identified as a more distantly related group of cyclins (Leopold and O'Farrell, 1991; Lew *et al.*, 1991): Cyclin H (THIIH), Cyclin C and Cyclin T (T for transcription). Despite their name, these cyclins do not show major fluctuation in abundance throughout the cell cycle, clearly separating therefore the regulation and role of the CTD kinases from classical cell cycle CDKs (Adamczewski *et al.*, 1996; Garriga *et al.*, 1998; Nigg, 1996; Rickert *et al.*, 1996; Tassan *et al.*, 1994). Nevertheless, there are many clues to a link between

transcription and the cell cycle through these kinases (Edwards *et al.*, 1998; Fu *et al.*, 1999; Nishiwaki *et al.*, 2000; Ren and Rollins, 2004; Sage, 2004).

2.2. Three CDKs differentially phosphorylate the CTD and regulate transcription

2.2.1. CDK7/Cyclin H

The specificity of CDK7/Cyclin H for the CTD seems to depend on the association with higher order complexes. The CDK/cyclin pair for itself and in complex with a third protein – *ménage à trois* –, the MAT1 RING finger protein forms the so-called CAK, the CDK activating kinase, which phosphorylates other CDKs and thereby activates them during cell cycle (Devault *et al.*, 1995; Fisher *et al.*, 1995; Harper and Elledge, 1998; Larochelle *et al.*, 1998; Wallenfang and Seydoux, 2002). The affinity towards the CTD is increased when this trimeric complex is part of the general transcription factor TFIIH (Rossignol *et al.*, 1997; Yankulov and Bentley, 1997) and the substrate specificity towards the CTD phosphorylation is highest when TFIIH is part of the RNAPII and GTF containing initiation complex (Lu *et al.*, 1992; Watanabe *et al.*, 2000). Phosphorylation of the CTD by CDK7 during initiation leads to promoter escape. Kin28 has been proposed to be the primary CTD kinase at initiation *in vivo* (Komarnitsky *et al.*, 2000; Liu *et al.*, 2004; Schroeder *et al.*, 2000; Valay *et al.*, 1995). TFIIH has at least three functions in transcription: ATP-dependent promoter DNA opening by its helicase activities, CTD phosphorylation and suppression of arrest of early RNAPII elongation complexes (Oelgeschlager, 2002). In *S. cerevisiae* the CDK7 ortholog Kin28 is apparently less linked to the cell cycle since it exhibits no CAK activity, which instead is mediated by a different CDK activating kinase, Cak1 (Kaldis *et al.*, 1996). *S. pombe* possesses two CAKs (Hermand *et al.*, 1998; Lee *et al.*, 1999a; Molz and Beach, 1993). Kin28 is the kinase present in TFIIH in *S. cerevisiae* and is required for normal levels of transcripts *in vivo* (Holstege *et al.*, 1998). The kinase activity of Kin28 is essential for viability (Liu *et al.*, 2004).

2.2.2. CDK8/Cyclin C

The CDK8/Cyclin C pair (or Srb10/Srb11) was discovered in yeast and human cells (Liao *et al.*, 1995; Tassan *et al.*, 1995). As described above CDK8/Cyclin C

associates with MED12 (Srb8) and MED13 (Srb9), to form a fourth module of the Mediator of transcriptional regulation, the CDK8/Cyclin C module. This module is present in a subpopulation of the Mediator, is conserved among eukaryotes, and phosphorylates the Pol II CTD (Borggreffe *et al.*, 2002; Boubé *et al.*, 2002; Liu *et al.*, 2001; Samuelson *et al.*, 2003). As CDK7, CDK8 is found tightly associated with the preinitiation complex (Liu *et al.*, 2001; Orphanides *et al.*, 1996; Ranish *et al.*, 1999). The first identification of this kinase had already linked it to RNAPII transcription and the CTD. It was cloned together with the mentioned *srb* genes as suppressors of a temperature sensitive phenotype caused by truncations of the CTD (Koleske *et al.*, 1992; Koleske and Young, 1994; Thompson *et al.*, 1993) and turned out to form large parts of the Mediator complex. The CDK8/Cyclin C module is a target of the Ras/PKA signal transduction pathway (Chang *et al.*, 2004) and is moreover involved in Notch signaling (Fryer *et al.*, 2004). The role of phosphorylation by CDK8 in transcription still remains enigmatic. An appealing model is the function as a repressor of transcriptional initiation: formation of the preinitiation complex could be prevented through premature phosphorylation of the CTD (Hengartner *et al.*, 1998). On the other hand a recent study points out an overlapping function of CDK7 and CDK8 *in vivo*, with the essential CDK7 being the more dominant kinase masking *in vivo* effects of a CDK8 deletion (Liu *et al.*, 2004). Upon deletion of this kinase several genes are derepressed and many others are repressed, but it is still unclear whether this effect of repression is direct or indirect. There are also many other potential substrates involved in transcription that are targets of CDK8, such as activators, like Gal4 and Sip4 (Ansari *et al.*, 2002; Vincent *et al.*, 2001), and the factors Gcn4 and Ste12, which are destabilized upon phosphorylation (Chi *et al.*, 2001; Nelson *et al.*, 2003). CDK8 binds the Gal4 activation domain (Ansari *et al.*, 2002). CDK8 promotes ATP-dependent dissociation of preinitiation complexes, resulting in a positive effect on transcription (Liu *et al.*, 2004). CDK8/Cyclin C phosphorylates subunits of the general transcription factor TFIID (Liu *et al.*, 2004), and the Mediator subunit MED2 (Hallberg *et al.*, 2004). Mutation of the MED2 phosphorylation site reduces expression levels of certain genes (Hallberg *et al.*, 2004). The role of the binding partners within the CDK8/Cyclin C module for recruiting the kinase and thereby influencing substrate specificity remains to be eluted.

As for CDK7, also for CDK8/CyclinC an additional role in the cell cycle has been implicated. Human CDK8 is able to inactivate CDK7 through phosphorylation

and might therefore repress its CDK activating activity as well as transcription (Akoulitchev *et al.*, 2000). The regulatory Cyclin C, which had originally been identified to perform a role in G1/S transition (Leopold and O'Farrell, 1991; Lew *et al.*, 1991), was recently shown to play a role in the cell cycle by association with CDK3 in human cells (Ren and Rollins, 2004; Sage, 2004). Cyclin C could also play a second role in transcription through association with CDK11, a recently characterized kinase that coordinates transcription and RNA processing (Hu *et al.*, 2003; Trembley *et al.*, 2003).

2.2.3. CDK9/Cyclin T

P-TEFb had originally been identified in metazoans as a heterodimer of CDK9 with one of the cyclin T isoforms T1, T2, or K (Peng *et al.*, 1998). A larger P-TEFb complex with reduced activity additionally contains the small nuclear RNA 7SK and the HEXIM protein (Michels *et al.*, 2003; Nguyen *et al.*, 2001; Yang *et al.*, 2001; Yik *et al.*, 2003). P-TEFb was isolated by the ability to overcome arrest by RNAPII complexes in early elongation (Marshall and Price, 1995) and its role in phosphorylating the CTD is implied by the fact that *in vitro* this only occurs when the CTD is present on RNAPII (Marshall *et al.*, 1996).

P-TEFb is thought to function in a dual way through its effect on 5,6-dichloro-1-beta-d-ribofuranosylbenzimidazole (DRB) sensitivity inducing factors (DSIF) and the negative elongation factor (NELF), which repress early elongation by interacting with hypophosphorylated RNA Pol II. On one hand P-TEFb could act through transition from the hypophosphorylated to the hyperphosphorylated RNA Pol II form and on the other hand directly act upon DSIF, which can be phosphorylated by CDK9 on its Spt5 subunit (Wada *et al.*, 1998a; Wada *et al.*, 1998b; Yamaguchi *et al.*, 1999a; Yamaguchi *et al.*, 1998; Yamaguchi *et al.*, 1999b).

CTD kinase 1 (CTDK1) and Bur1/Bur2 are the two closely related CDKs that could in *S. cerevisiae* comprise a similar function as P-TEFb (Guo and Stiller, 2004; Lee and Greenleaf, 1991; Murray *et al.*, 2001; Prelich and Winston, 1993; Sterner *et al.*, 1995). While Bur1/Bur2 represent a CDK and its cyclin, CTDK1 consists, similar to the other CTD kinases, of an additional subunit: it is composed of three subunits: Ctk1, Ctk2 and Ctk3 (Carboxy Terminal Domain kinase). Bur1 was suggested to be the specific ortholog of CDK9 after computational analysis (Guo and Stiller, 2004). Chromatin immunoprecipitation (ChIP) and genetic data suggest that the two kinases

do not overlap in their function but both play a crucial role in elongation (Cho *et al.*, 2001; Keogh *et al.*, 2003; Yao *et al.*, 2000; Yao and Prelich, 2002). While Bur1 is essential Ctk1 is not (Lee and Greenleaf, 1991; Prelich, 2002). It was recently suggested (Keogh *et al.*, 2003) that Ctk1 might have the CTD as a substrate, while Bur1 phosphorylates Spt4/5: Bur1/2 share a number of phenotypes with *spt4/5/6* deletions (Happel *et al.*, 1991; Neigeborn *et al.*, 1987; Prelich and Winston, 1993; Swanson *et al.*, 1991; Swanson and Winston, 1992) and Spt5 contains several sequence motifs reminiscent of the CTD heptapeptide (Keogh *et al.*, 2003). This model is particularly appealing since the situation would be similar to another split kinase function in yeast: *S. cerevisiae* also has two kinases for CAK activity and CTD phosphorylation by TFIIH, instead of one, as for example in the human system.

2.2.4. Substrate specificity and regulation of CTD kinases

Phosphorylation occurs mainly at Ser2 and Ser5 (Corden *et al.*, 1985) with around 50 phosphates groups in the hyperphosphorylated RNA Pol II form, an average of one per repeat (Payne and Dahmus, 1993). More Ser5 phosphorylation has been observed in promoter proximal regions, Ser2 phosphorylation predominates in more distal regions and triggers binding of the 3'-RNA processing machinery (Cho *et al.*, 2001; Komarnitsky *et al.*, 2000). Ser5 phosphorylation recruits and activates the mRNA capping machinery (Cho *et al.*, 1997; Ho *et al.*, 1998; Komarnitsky *et al.*, 2000; McCracken *et al.*, 1997). The CDKs involved in CTD phosphorylation differ in their specificity towards the target serine residue. CDK7 and CDK8 phosphorylate S5, whereas CDK9/Ctk1/Bur1 phosphorylates S2, although depending on the assay used, the length of the peptide, and detection mechanism different results have been obtained (Table 2). Thus CDK7 activity predominates during initiation, whereas CDK9 activity is important during elongation (Cho *et al.*, 2001; Kim *et al.*, 2002; Komarnitsky *et al.*, 2000). CDK7 and CDK8, but not Bur1 or Ctk1, are stably associated with transcription initiation complexes (Liu *et al.*, 2004). CDKs also differ in their processivity. CDK7 generates hyperphosphorylated CTD peptides, whereas CDK8 and CDK9 generate CTD peptides with fewer phosphorylations (Pinhero *et al.*, 2004). A few principles underlying CTD phosphorylation state recognition have been suggested (Meinhart and Cramer, 2004; Meinhart *et al.*, 2005).

Kinase specificity for the CTD may not only be achieved by CTD recognition at the kinase active site, but also by CTD binding to kinase-associated factors. The preference of TFIIF for S5 phosphorylation is enforced by TFIIE (Yamamoto *et al.*, 2001). The HIV Tat protein shifts CDK9 phosphorylation preference from S2 to both S2 and S5 (Zhou *et al.*, 2000). Cyclin A has a conserved surface patch that binds kinase substrates (Kontopidis *et al.*, 2003; Schulman *et al.*, 1998). Cyclin T binds the CTD via a histidine-rich stretch in its C-terminal domain (Kurosu *et al.*, 2004; Taube *et al.*, 2002). A recent study suggests that the cyclins generally act as adaptors to render a CDK specific for a substrate (Loog and Morgan, 2005).

Table 2 – Specificity of CTD kinases

CTD kinase	S2 phosphorylation	S5 phosphorylation	Reference
CDK7		+	(Rickert <i>et al.</i> , 1999)
		+	(Trigon <i>et al.</i> , 1998)
	(+)		(Dubois <i>et al.</i> , 1997; Patturajan <i>et al.</i> , 1998)
		+	(Zhou <i>et al.</i> , 2000)
		+	(Ramanathan <i>et al.</i> , 2001)
		+	(Sun <i>et al.</i> , 1998)
	+	+	(Watanabe <i>et al.</i> , 2000; Yamamoto <i>et al.</i> , 2001)
		+	(Kim <i>et al.</i> , 2002)
	(+)	+	(Komarnitsky <i>et al.</i> , 2000)
		+	(Hengartner <i>et al.</i> , 1998)
CDK8	+	+	(Sun <i>et al.</i> , 1998)
		+	(Rickert <i>et al.</i> , 1999)
		+	(Ramanathan <i>et al.</i> , 2001)
		+	(Hengartner <i>et al.</i> , 1998)
CDK9	+	+	(Borggreffe <i>et al.</i> , 2002)
	+	(+)	(Komarnitsky <i>et al.</i> , 2000)
	+	(+)	(Zhou <i>et al.</i> , 2000)
		+	(Ramanathan <i>et al.</i> , 2001; Ramanathan <i>et al.</i> , 1999)
	+		(Shim <i>et al.</i> , 2002)
Ctk1	+		(Kim <i>et al.</i> , 2002)
	+		(Dubois <i>et al.</i> , 1997; Patturajan <i>et al.</i> , 1998)
	+		(Cho <i>et al.</i> , 2001)
	+	+	(Jones <i>et al.</i> , 2004)
Bur1		+	(Murray <i>et al.</i> , 2001)

An open question is the activation mechanism of the CTD-targeting CDKs. CDKs involved in cell cycle regulation are generally activated in two steps, cyclin binding, and phosphorylation of a conserved threonine in the CDK activation segment (T160 in human CDK2) (Pavletich, 1999). Interaction of the phosphothreonine side chain with three conserved arginines triggers a conformational change that results in full kinase activation (Russo *et al.*, 1996). CDK7 and CDK9 carry a threonine or a serine at the phosphorylated position. In the free CDK7 structure, the phosphorylated threonine however is found at a different location than in CDK2 (Lolli *et al.*, 2004), and does not contact the three conserved arginines, pointing to a different mechanism of CDK activation. Also, CDK8 does not have a threonine or serine residue at the position phosphorylated in other CDKs (Tassan *et al.*, 1995).

Structural information about the CTD kinases, compared to other CDKs and cyclins, helps to shed light on their regulation mechanisms. The structure of Cyclin H revealed a difference to other known cyclin structures (Andersen *et al.*, 1996a; Lolli *et al.*, 2004). Its N- and C-terminal helices form an additional domain, which the authors speculate to be strictly required for structural integrity of the protein. These results are consistent with the observation that the N-terminal helix in Cyclin H is held in place by highly conserved and at the same time specific interactions with the cyclin fold that do not allow for much flexibility. In other cyclins that regulate cell cycle progression the N-terminal helices were seen involved in CDK binding (Card *et al.*, 2000; Jeffrey *et al.*, 1995). The MAT1 protein was proposed to bind to CDK7 itself rather the cyclin subunit since MAT1 interacts specifically and directly with CDK7 independently of the presence or absence of Cyclin H (Andersen *et al.*, 1996a; Lolli *et al.*, 2004).

3. Aims of this work

To elucidate the enigmatic molecular mechanism of Mediator it is necessary to answer questions using structural information. When this work was started, detailed structural information on Mediator had not been available. The reasons for this lie in the difficulties to obtain homogenous and pure fractions from native Mediator preparations caused by the modularity and low abundance of Mediator in the cell. Recombinant overexpression had so far not been successful in amounts suitable for structural studies since multiple protein-protein interactions in the natural complex do not allow soluble expression of individual subunits.

One main objective of this work was to define an interaction map and stable subcomplexes within the Mediator head module and to establish a technique that allowed high expression levels of recombinant subcomplexes. These subcomplexes can be used for biochemical assays, establishing function towards RNA Pol II, e.g. determining the RNA Pol II interacting subunits, and can be used for structural studies with and without RNA Pol II using crystallography. As a short-term goal this approach can lead to high-resolution structural information of subcomplexes and subunits within Mediator. In the future, these structures can be located in lower resolution maps of Mediator becoming available with and without Pol II through crystallographic and EM techniques. A recombinant core Mediator produced by overexpression could in turn lead to higher resolutions in structural studies since it allows absolute homogeneity and can be used for biochemical studies to shed light on its mechanism in activating RNA Pol II.

The second main objective of this work was to obtain structural information on the Mediator CDK8/Cyclin C pair. The CDK8/Cyclin C module of Mediator comprises a kinase activity and is involved in transcriptional regulation. At the time this work was started the only structural information on CDK/cyclin pairs involved in transcription was limited to the crystal structure of Cyclin H (Andersen *et al.*, 1996a; Lolli *et al.*, 2004). It had revealed differences in fold when compared to other cyclins active in the cell cycle. Detailed structural information on the CDK8/Cyclin C heterodimer would provide further understanding of such differences. With the structure of Cyclin C the importance of this cyclin in complex formation and substrate specificity can be understood. Three dimensional (3D) modeling techniques have the potential to give first understanding on the specific features of the Mediator kinase

and allow for speculations on its phosphorylation independent activation as well as specific and non-specific recognition of CDK and cyclin.

Materials and Methods

1. Bacterial strains and insect cells

Bacterial strains and insect cell lines used in this work are listed in table 3 and 4 respectively.

Table 3 – bacterial strains

Strain	Genotype or description	Source or reference
<i>Escherichia coli</i>		
DH5 α	F' Φ 80dlacZ Δ M15 Δ (lacZYA-argF)U169 deoR recA1 endA1 hsdR17(r _K ⁻ m _K ⁺) phoA supE44 λ - thi-1 gyrA96 relA1	(Woodcock <i>et al.</i> , 1989)
XL-1 blue strain	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac[F' proAB lac ^f Z Δ M15Tn10(Tet ^f)]	Stratagene
BL21-CodonPlus (DE3)-RIL	B F ⁻ ompT hsdS(r _B ⁻ m _B ⁻) dcm ⁺ Tet ^r gal ₋ (DE3) endA Hte [argU ileY leuW Cam ^r]	Stratagene
DH10-Bac TM	F- mcrA Δ (mrr-hsdRMS-mcrBC) ₋₈₀ lacZ Δ M15 Δ lacX74 recA1 endA1 araD139 Δ (ara, leu)7697 galU galk rpsL nupG/bMON14272/pMON7124	Invitrogen
BL21-(DE3)		Stratagene
B834	<i>E. coli</i> (DE3) (hsd metB)	(Budisa <i>et al.</i> , 1995)

Table 4 – insect cells

Cell line	Description	Source or reference
<i>Spodoptera frugiperda</i> SF9	Cell line from <i>S. frugiperda</i> gonad cells	Novagen

2. Plasmids and genomic DNA

Plasmids constructed and used for this work are listed in table 5. Genomic DNA from *S. cerevisiae* and cDNA from *S. pombe* was used as template for PCR. Names refer to *S. cerevisiae* genes unless marked by *Sp* (*S. pombe*). His₆ - tags were always introduced on the C-terminal end.

Table 5 – plasmids

Nº	Name	Vector	AB ^R	Insert	Restriction sites	Affinity tag	Remarks
1	Srb4 (MED17)	pGex-3x	amp	MED17	<i>Bam</i> Hl	N-term GST	
2	Srb4-F2 (MED17-F2)	pGex-3x	amp	MED17 ₂₄₁₋₆₈₇	<i>Bam</i> Hl	N-term GST	
3	Srb4 (MED17)	pET21b	amp	MED17	<i>Nhe</i> l/ <i>Not</i> I	none	Stop codon 5' from His ₆
4	Srb4ΔC (MED17ΔC)	pET21b	amp	MED17 ₁₋₆₃₀	<i>Nhe</i> l/ <i>Not</i> I	His ₆	
5	Srb4ΔN (MED17ΔN)	pET21b	amp	MED17 ₉₇₋₆₈₇	<i>Nhe</i> l/ <i>Not</i> I	His ₆	
6	Srb4-ΔCΔN (MED17ΔCΔN)	pET21b	amp	MED17 ₉₇₋₆₃₀	<i>Nhe</i> l/ <i>Not</i> I	His ₆	
7	Srb4-F1 (MED17-F1)	pET21b	amp	MED17 ₁₋₂₄₀	<i>Nhe</i> l/ <i>Not</i> I	His ₆	
8	Srb4-D2 (MED17-D2)	pET21b	amp	MED17 ₁₋₃₅₁	<i>Nhe</i> l/ <i>Not</i> I	His ₆	
9	Srb4-D2ΔN (MED17-D2ΔN)	pET21b	amp	MED17 ₉₇₋₃₅₁	<i>Nhe</i> l/ <i>Not</i> I	His ₆	
10	Srb4-F2-His	pET21b	amp	MED17 ₂₄₁₋₆₈₇	<i>Nde</i> l/ <i>Not</i> I	none	Stop codon before C-term His ₆
11	Srb4-F2ΔC (MED17-F2ΔC)	pET21b	amp	MED17 ₂₄₁₋₆₃₀	<i>Nde</i> l/ <i>Not</i> I	His ₆	
12	Rox3-Srb7	pET21b	amp	MED19 MED21	<i>Nhe</i> l/ <i>Eco</i> RI <i>Nde</i> l/ <i>Not</i> I	- His ₆	Bicistron I
13	Srb2-Srb7	pET21b	amp	MED20 MED21	<i>Nhe</i> l/ <i>Eco</i> RI <i>Nde</i> l/ <i>Not</i> I	- His ₆	Bicistron I
14	Med11-Srb7	pET21b	amp	MED11 MED21	<i>Nhe</i> l/ <i>Eco</i> RI <i>Nde</i> l/ <i>Not</i> I	- His ₆	Bicistron I
15	Med8-Srb7	pET21b	amp	MED8 MED21	<i>Nhe</i> l/ <i>Eco</i> RI <i>Nde</i> l/ <i>Not</i> I	- His ₆	Bicistron I
16	Med7-MED17-F2	pET24d	kan	MED7GSTtev MED17-F2 ₂₄₁₋₆₈₇	<i>Nhe</i> l/ <i>Eco</i> RI <i>Nde</i> l/ <i>Not</i> I	- His ₆	Bicistron I
17	Rox3-Srb4-F2	pET21b	amp	MED19 MED17 ₂₄₁₋₆₈₇	<i>Nhe</i> l/ <i>Eco</i> RI <i>Nde</i> l/ <i>Not</i> I	- His ₆	Bicistron I
18	Srb2-Srb4-F2	pET21b	amp	MED20 MED17 ₂₄₁₋₆₈₇	<i>Nhe</i> l/ <i>Eco</i> RI <i>Nde</i> l/ <i>Not</i> I	- His ₆	Bicistron I
19	Med11-Srb4-F2	pET21b	amp	MED11 MED17 ₂₄₁₋₆₈₇	<i>Nhe</i> l/ <i>Eco</i> RI <i>Nde</i> l/ <i>Not</i> I	- His ₆	Bicistron I
20	Med8-Srb4-F2	pET21b	amp	MED8 MED17 ₂₄₁₋₆₈₇	<i>Nhe</i> l/ <i>Eco</i> RI <i>Nde</i> l/ <i>Not</i> I	- His ₆	Bicistron I
21	Srb6-Srb4-F2	pET21b	amp	MED22 MED17 ₂₄₁₋₆₈₇	<i>Nhe</i> l/ <i>Eco</i> RI <i>Nde</i> l/ <i>Not</i> I	- His ₆	Bicistron I
22	Srb6-Srb4-F2	pET21b	amp	MED22 MED17 ₂₄₁₋₆₈₇	<i>Nhe</i> l/ <i>Eco</i> RI <i>Nde</i> l/ <i>Not</i> I	- His ₆	Bicistron I
23	Med6ΔC-Srb4-F2	pET21b	amp	MED6 ₁₋₂₁₄ MED17 ₂₄₁₋₆₈₇	<i>Nhe</i> l/ <i>Eco</i> RI <i>Nde</i> l/ <i>Not</i> I	- His ₆	Bicistron I
24	Srb4-F1ΔN (MED17-F1ΔN)	pET21b	amp	MED17 ₉₇₋₂₄₀	<i>Nhe</i> l/ <i>Not</i> I	His ₆	

31	Srb10-Srb11 _{Sp}	pET21b	amp	CDK8 _{Sp} CycC _{Sp}	<i>NheI/EcoRI</i> <i>NcoI/NotI</i>	- His ₆	Bicistron II
32	Srb10 _{Sp}	pET21b	amp	CDK8 _{Sp}	<i>NheI/NotI</i>	His ₆	
33	Srb11 _{Sp}	pET28d	kan	CycC _{Sp}	<i>NcoI/NotI</i>	His ₆	
34	Srb10 _{Sp} -Bac	pFastBac1	amp	CDK8 _{Sp}	<i>BSSHI/ EcoRI</i>	His ₆	Invitrogen, His Tag included in primer
35	Srb2-Srb5	pET21b	amp	MED20 MED18	<i>NheI/EcoRI</i> <i>NcoI/NotI</i>	- His ₆	Bicistron II
36	Srb11 _{Sp} -1	pET28d	kan	CycC _{Sp} 5-228	<i>NcoI/NotI</i>	none	
37	Srb11 _{Sp} -2	pET28d	kan	CycC _{Sp} 30-228	<i>NcoI/NotI</i>	none	
38	Srb11 _{Sp} -3	pET28d	kan	CycC _{Sp} 5-228	<i>NcoI/NotI</i>	His ₆	
39	Srb11 _{Sp} -4	pET28d	kan	CycC _{Sp} 30-228	<i>NcoI/NotI</i>	His ₆	
40	Srb11 _{Sp} -Meth1	pET28d	kan	CycC _{Sp} L60M	<i>NcoI/NotI</i>	His ₆	
41	Srb11 _{Sp} -Meth2	pET28d	kan	CycC _{Sp} I125M	<i>NcoI/NotI</i>	His ₆	
42	Srb2-Srb5	pET24b	kan	MED20 MED18	<i>NheI/EcoRI</i> <i>NcoI/NotI</i>	- His ₆	Bicistron II
43	Srb11-3 _{Sp} - Meth1	pET28d	kan	CycC _{Sp} 5-228 L60M	<i>NcoI/NotI</i>	His ₆	
44	Srb11-3 _{Sp} - Meth2	pET28d	kan	CycC _{Sp} 5-228 I125M	<i>NcoI/NotI</i>	His ₆	
45	Srb11-3 _{Sp} - Meth3	pET28d	kan	CycC _{Sp} 5-228 L186M	<i>NcoI/NotI</i>	His ₆	
46	Srb11-3 _{Sp} - Meth1,3	pET28d	kan	CycC _{Sp} 5-228 L60M, L186M	<i>NcoI/NotI</i>	His ₆	
47	Srb11-3 _{Sp} - Meth1,2,3	pET28d	kan	CycC _{Sp} 5-228 L60M, I125M, L186M	<i>NcoI/NotI</i>	His ₆	
48	Srb2 _{Sp} -Srb2- Srb5	pET24b	kan	MED20 _{Sp} MED20 MED18	<i>KpnI/SacI</i> <i>NheI/EcoRI</i> <i>NcoI/NotI</i>	- - His ₆	Tricistron II Two <i>SacI</i> sites
49	Srb2 _{Sp} -Srb5- ΔC/N	pET24b	kan	MED20 _{Sp} - MED18 ₅₆₋₂₈₅	<i>KpnI/SacI</i> - <i>BamHI/NotI</i>	- - His ₆	Bicistron IV
50	Srb2 _{Sp} -Srb2- Srb5ΔC/N	pET24b	kan	MED20 _{Sp} MED20 MED18 ₅₆₋₂₈₅	<i>KpnI/SacI</i> <i>NheI/EcoRI</i> <i>BamHI/NotI</i>	- - His ₆	Tricistron III
51	Med8-Srb2- Srb5ΔC/N	pET24b	kan	MED8 MED20 MED18 ₅₆₋₂₈₅	<i>NdeI/SacII</i> <i>NheI/EcoRI</i> <i>BamHI/NotI</i>	- - His ₆	Tricistron IV
52	Srb2 _{Sp} -Srb2- Srb5ΔC/N	pET24b	kan	MED20 _{Sp} MED20 MED18 ₅₆₋₂₈₅	<i>KpnI/HindIII</i> <i>NheI/EcoRI</i> <i>BamHI/NotI</i>	- - His ₆	Tricistron V
53	Srb2 _{Sp} -Srb5 _{Sp} - Srb5ΔC/N	pET24b	kan	MED20 _{Sp} MED18 _{Sp} MED18 ₅₆₋₂₈₅	<i>KpnI/HindIII</i> <i>NheI/SacI</i> <i>BamHI/NotI</i>	- - His ₆	Tricistron VI
54	Med8-Srb2- Srb5	pET24b	kan	MED8 MED20 MED18	<i>NheI/EcoRI</i> <i>NheI/EcoRI</i> <i>NcoI/NotI</i>	- - His ₆	Tricistron I
55	Srb2-Srb5ΔC/N	pET24b	kan	MED20 MED18 ₅₆₋₂₈₅	<i>NheI/EcoRI</i> <i>BamHI/NotI</i>	- His ₆	Bicistron III
56	Srb2-Srb5-ΔN1	pET24b	kan	MED20 MED18 ₁₄₁₋₃₀₇	<i>NheI/EcoRI</i> <i>NcoI/NotI</i>	- His ₆	Bicistron II
57	Srb4-F2- MED6ΔC	pGex-3x	amp	MED17 ₂₄₁₋₆₈₇ MED6 ₁₋₂₁₄	<i>BamHI</i> <i>NdeI/XmaI</i>	N-term GST His ₆	Claudia Buchen

58	Srb4-F2-MED6	pGex-3x	amp	MED17 ₂₄₁₋₆₈₇ MED6	<i>BamHI</i> <i>NdeI/XmaI</i>	N-term GST His ₆	Claudia Buchen
59	MED6ΔC	pET24d	kan	MED6 ₁₋₂₁₄	<i>NheI/EcoRI</i>	His ₆	
60	pfastBacHT A	pFastBacHT _A	amp	None	MCS	N-term His ₆	Invitrogen
61	pfastBacHT B	pFastBacHT _B	amp	None	MCS	N-term His ₆	Invitrogen
62	pFastBacHT C	pFastBacHT _C	amp	None	MCS	N-term His ₆	Invitrogen
63	pFastBac1	pFastBac	amp	None	MCS	none	Invitrogen
64	pFastBac Dual	pFastBac Dual	amp	None	MCS	none	Invitrogen
65	RY7488	pSK157	amp	CDK8	<i>XbaI/BamHI</i>	none	R. Young –baculovirus transfer vector
66	RY7489	pSK158	amp	CDK8	<i>XbaI/BamHI</i>	HA	“
67	RY7490	pSK159	amp	Cyclin C	<i>EcoRI/BamHI</i>	none	“
68	RY7491	psK160	amp	Cyclin C	<i>EcoRI/BamHI</i>	HA	“
69	Srb10-Bacmid	Invitrogen Bacmid from DH10-Bac™ cells	kan	CDK8	Transposon target sites	His ₆	For reamplification use glycerol stock

Antibiotic resistences (AB^R) are abbreviated as amp (ampicillin) and kan (kanamycin).

3. Media

Growth media that were used for recombinant expression and the according supplements and antibiotics are listed in tables 6 and 7 respectively.

Table 6 – growth media

LB	(Sambrook and Russel, 2001)	1% tryptone; 0.5% yeast extract; 0.5% NaCl
SOC	(Sambrook and Russel, 2001)	2.0% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl ₂ , 20 mM MgSO ₄ , 20 mM glucose
SF-900 II	(PAA)	Serum free medium for insect cell culture (recipe see PAA catalogue)

Table 7 – supplements and antibiotics

Supplements and Antibiotics	stock solution	Applied to strain or cell line as follows		
		<i>E. coli</i>	<i>E. coli</i> DH10Bac	SF9 insect cells
Ampicillin	100 mg/ml in H ₂ O	100 µg/ml	--	--
Gentamycin	50 mg/ml in H ₂ O	--	7 µg/ml	--
Kanamycin	30 mg/ml in H ₂ O	30 µg/ml	50 µg/ml	--
Tetracyclin	12.5 mg/ml	--	10 µg/ml	--
Chloramphenicol	50 mg/ml in ethanol	50 µg/ml	20 µg/ml	--
IPTG	0.5 M in H ₂ O	0.5 mM	40 µg/ml	--
X-Gal	40 mg/ml in dimethyl formamide	--	100 µg/ml	--
Baculo Gold FCS	100%	--	--	0-10%
Penicillin/ Streptomycin 100x solution (PAA)	1000 units/ml pen and 1000 µg/ml strep	--	--	50 units/ml pen; 50µg/ml strep
DMSO (cell culture grade)	100%	--	--	10%

Baculo FCS Gold was heat inactivated for 30 min at 56° before aliquotation in 50 ml.

4. Buffers and solutions

Tables 8-14 list the buffers and solutions used in this work. Buffers for individual protein purifications are listed separately along with purification protocols.

Table 8 – buffers and solutions used SDS-PAGE

4x stacking gel buffer	1.5 M Tris; 0.4% (w/v) SDS; pH 8.8
4x separation gel buffer	1.5 M Tris; 0.4% (w/v) SDS; pH 6.8
electrophoresis buffer	1,25 M glycine; 125 mM Tris/Cl; 0.5% (w/v) SDS
2x sample suffer	10% (v/v) glycerol, 50 mM Tris-HCl pH 7.0, 0.1% (w/v) bromophenol blue, 0.1% (w/v) lauryl sulfate, 1% (w/v) β-mercaptoethanol, 14% (w/v) 1,4-dithiothreitol

Table 9 – buffers and solutions for SDS- polyacrylamide gel staining

Coomassie stain solutions	
stain solution	50% (v/v) ethanol, 7% (v/v) acetic acid, 0.125% (w/v) Coomassie Brilliant Blue R-250
Destain solution	5% (v/v) methanol, 7.5% (v/v) acetic acid
Silver stain solutions (Bloom <i>et al.</i> , 1987)	
fixing solution	50% (v/v) methanol, 12.5% (v/v) acetic acid, 0.0001% (w/v) formaldehyde (for 50ml: 25 µl of 20% solution)
rinse solution I	50% (v/v) ethanol
rinse solution II	30% (v/v) ethanol
sensitizer	0.02% (w/v) (0.2 mg/ml) sodium thiosulphate
stain solution	0.1% (w/v) silver nitrate, 0.00015% (w/v) formaldehyde
developer	6% (w/v) sodium carbonate, 0.0001% (w/v) formaldehyde, 0.002% (v/v) sensitizer solution (100 µl sensitizer for 50 ml)
stop solution	5% (v/v) acetic acid

Table 10 – buffers and solutions for Western Blot

Transfer buffer	250 mM glycine; 25 mM Tris/HCl; 0,1% (w/v) SDS; 20% (v/v) methanol
Blotting buffer	0.1% (v/v) Tween 20 in PBS
Blocking buffer	0.1% (v/v) Tween 20, 5% (w/v) skimmed milk powder in PBS

Table 11 – buffers solutions for agarose Gels

6x Sample Buffer	1.5 mg/ml bromophenol blue, 1.5 mg/ml xylene cyanole, 50% (v/v) glycerol
1x TBE-Electrophoresis-buffer	89 mM Tris, 89 mM boric acid, 2 mM EDTA

Table 12 – buffers for the preparation of competent cells

TFB-1	30 mM KOAc, 50 mM MnCl ₂ , 100 mM RbCl, 10 mM CaCl ₂ , 15% (v/v) glycerol, pH 5.8
TFB-2	10 mM MOPS (pH 7.0), 10 mM RbCl, 75 mM CaCl ₂ , 15% (v/v) glycerol

Table 13 – sample preparation for Edman sequencing

Blotting buffer	10% (v/v) methanol in ddH ₂ O
Rehydration buffer	200 mM Tris/HCl pH 8.5, 2% (w/v) SDS

Table 14 – GST-CTD pulldown assays

Buffer A	20 mM Hepes pH 7.5, 150 mM NaCl, 10 mM DTT, 1 mM Pefabloc
Buffer B	20 mM Hepes pH 7.5, 1000 mM NaCl, 10 mM DTT, 1 mM Pefabloc

5. Molecular cloning techniques

5.1. Oligonucleotides

PCR primers for cloning genes were constructed in a standard way: after a 9 nt overhang to assure efficiency of cleavage the corresponding restriction sites were introduced followed by 20 to 25 nt of the sequence of interest on both sites of the gene. The primers for construction of the bicistron were designed according to (Lutzmann *et al.*, 2002).

5.2. PCR

Template for PCR was either genomic DNA from *S. cerevisiae* or cDNA from *S. pombe*. DNA was amplified either with Herculase (Fermentas) for molecular cloning or *Thermus aquaticus* (Taq) DNA polymerase for analytic PCR reactions with appropriate buffer systems from different commercial and non-commercial sources. Generally, 50 μl reactions contained 100 μM of each of the four dNTPs, 25 pmol of each primer, between 1 mM and 2 mM MgCl_2 (for Taq only) and approximately 1 ng of template DNA. For Herculase reactions DMSO was added according to the manufacturers suggestions. Thermocycling program was performed in 30 cycles. Times and temperatures of denaturation, annealing and elongation were moderately varied to meet the special requirements of the polymerase and primer-template pairs used in different amplifications.

5.3. DNA isolation

E. coli cells from 5 ml of an overnight culture were sedimented by centrifugation. Plasmid DNA was extracted from the pellet using the Spin Minipreparation Kit (Qiagen) following the manufacturers instructions.

5.4. Restriction cleavage and dephosphorylation

DNA was cleaved using restriction enzymes and buffers supplied by MBI Fermentas and New England Biolabs (NEB) with standard protocols. The reaction was incubated over night when working with PCR fragments in order to facilitate cleavage close to end. To avoid religation, cleaved vector DNA was treated with calf intestine alkaline phosphatase (CIAP, Fermentas).

5.5. Ligation

Ligation of DNA fragments was conducted in 20 µl total volumes at 16°C for two to 16 hours using T4 ligase (MBI Fermentas) and T4 ligase buffer (MBI Fermentas). In most cases linearized vector was incubated with a large (approx. tenfold) excess of insert to improve results. Positive clones were verified by restriction analysis and sequencing.

5.6. Mutagenesis

Generation of point mutations in genes was achieved by the PCR overlap extension method (Higuchi *et al.*, 1988) where two overlapping PCR products are produced carrying the desired mutation in the primer. The resulting products serve as template in a second PCR round.

5.7. Transformation of plasmid DNA

The transformation of plasmid DNA into chemically competent *E. coli* was performed with standard heat shock protocols. After thawing a 50 µl aliquot of competent cells 2 µl of ligation mixture or 1 µl of purified plasmid DNA were added. Cells were incubated on ice for 30 min, heat shocked at 42°C for 45 sec and incubated on ice for another two min. Then, 450 µl of LB medium were added to the cells and the mixture was incubated for 1 h in a 37°C shaker before plating.

5.8. Preparation of competent cells

200 ml LB media were inoculated with 5 ml of an over-night culture of the bacterial strain. The cells were grown at 37°C until an OD₆₀₀ of 0.4-0.55 was reached. After incubation on ice for 10 minutes the cells were centrifuged for 10 minutes at 5,000 rpm and 4°C. All following steps were performed at 4°C. The pellet was washed with 50 ml TFB 1 buffer and centrifuged again. This pellet was resuspended in 4 ml TFB 2 buffer, before aliquoting and plunging into liquid nitrogen.

6. Biochemical methods

6.1. Protein expression and purifications

6.1.1. Protein expression in *E. coli* cells

For the expression of proteins, transformed BL21 CodonPlus (DE3)-RIL cells were inoculated from a preculture at a 1:100 ratio, grown in LB medium supplemented with antibiotics at 37°C to an OD₆₀₀ of 0.5-0.6 and flash cooled before expression was induced with 0.5 mM IPTG, and carried out over night at 18°C. For selenomethionine incorporation, the Cyclin C mutant protein was expressed in the methionine auxotroph *E. coli* strain B834 (DE3) (Budisa *et al.*, 1995). Cells were grown in eight liter LB medium, supplemented with kan (30 mg/l) to and OD of 0.6, harvested by centrifugation (3000 g), and were resuspended in eight liter of selenomethionine-containing minimal medium (Budisa *et al.*, 1995; Meinhart *et al.*, 2003). The cell suspension was agitated until growth resumed (2-3 hours) at 18°C, before expression was induced with 0.5 mM IPTG, and continued over night.

6.1.2. Cell lysis and chromatography

6.1.2.1. Cell lysis and affinity chromatography

Cells from a 2 l expression culture were lysed in a French Pressure Cell in 37 ml lysis buffer. The lysate was cleared by centrifugation (30 min, 16000 rpm, SS34 rotor) and applied to pre-equilibrated Ni-NTA agarose (1.5ml; Qiagen) or glutathione-sepharose (1 ml, Amersham) self assembled columns, depending on the affinity tag. The column was washed with 20 CV of lysis buffer and proteins were eluted with lysis buffer containing the corresponding competitive reagent. The elution step was optimized by gradient and varying pH for best yields and purities. For

column regeneration an elution step with 2 M imidazole or 100 mM glutathione elution buffer was carried out.

6.1.2.2. Ion exchange chromatography

Mono S HR 5/5, Mono Q HR 5/5, and Mono Q 10/100 GL (all Amersham) were used for ion exchange chromatography on the Äkta purifier system (Amersham). Before injection samples were diluted to the gradient starting conditions by adding low salt buffer and filtered. On the small columns the eluting salt gradient was usually applied over 15 CV, on the Mono Q 10/100 GL over 7 CV, ranging for example from 100 mM to 1000 mM NaCl.

6.1.2.3. Gel filtration

Samples generated by ion exchange chromatography were loaded onto equilibrated Superose6 10/30 HR or Superose12 10/30 HR gel filtration columns (Amersham). A maximum sample size of 500 µl was achieved by concentration of the samples. The gelfiltration was continued over at least 1 CV (24 ml) at maximum flow rates of 0.5 ml/min. To estimate the MW of proteins and complexes, the columns were calibrated with the Low and the High Molecular Weight Gel Filtration Calibration Kit (Amersham Pharmacia Biotech), which use proteins in the range between 13.7 kD and 669 kD.

6.1.3. Enrichment of proteins

In order to concentrate protein samples after intermediate and final purification steps centrifugal filter devices (Amicon® Ultra, Millipore or membra-spin Mini, membraPure) with 10 kD nominal molecular weight limit were used as described in the provided protocol. Concentration was conducted up to the desired volume for intermediate purification and up to the desired protein concentration in final concentration steps.

6.1.4. Individual purification protocols

In the following the individual purification protocols for proteins and multiprotein complexes are listed. Buffers are described as percentage of buffer B/C in buffer A. For 100x protease inhibitor mix 28.4 µg/ml leupeptin, 137 µg/ml pepstatin

A, 17 µg/ml methylbenzenesulfonyl chloride and 33 µg/ml benzamidine were dissolved in ethanol. For elution from affinity columns imidazole was added from a 2M stock pH 8.0 and glutathione was freshly dissolved in water and stock solution was brought to the corresponding pH by titration.

6.1.4.1. Purification of GST-MED17 (GST-Srb4)

Table 15 – buffers for GST-MED17 purification

Buffer A	50 mM Tris/HCl pH 8.0, 2 mM DTT
Buffer B	1000 mM NaCl, 50 mM Tris/HCl pH 8.0, 2 mM DTT
Elution	buffer A, 10 mM glutathione
Cleavage buffer	10% buffer B, 1 mM CaCl ₂ , 20 µl factor Xa

Typically bacterial from 2-8 l expressions were lysed by sonication in 15% buffer B containing protease inhibitors (Table 15). Cleared lysate was loaded on 1ml glutathione sepharose column, after washing with lysis buffer followed by buffer A protein was eluted as three 1 ml fractions. Pooled fractions were subsequently cleaved with factor Xa. Alternatively protease factor Xa reaction was performed on the column over night at 4° by applying 1 ml lysis buffer containing 20 U of factor Xa and 1 mM MgCl₂. Gelfiltration was performed with Superose12 and different salt concentrations (100-300 mM NaCl). After cleavage, protein could be used for limited proteolysis assays.

6.1.4.2. Purification of GST-MED17-F2 (GST-Srb4-F2)

Table 16 – buffers for GST-MED17-F2 purification

Buffer A	50 mM Tris/HCl pH 8.0, 10% (v/v) glycerol, 2 mM DTT
Buffer B	1000 mM NaCl, 50 mM Tris/HCl, 10% (v/v) glycerol, 2 mM DTT
Elution	30% buffer B, 10 mM glutathione
Cleavage buffer	Elution buffer, 1 mM CaCl ₂ , 30 µl factor Xa
Mono Q 5/5 HR	20-100% buffer B
Superose12	50% buffer B

Cells from 2 l were lysed in 30% buffer B containing protease inhibitors by sonication and cleared lysate was applied to a 1ml glutathione sepharose column (Table 16). After elution in 3 ml, CaCl₂ and factor Xa were added to perform cleavage over night at 4°C. The sample was diluted to meet anion exchange starting conditions. A gelfiltration was performed at relatively high salt conditions to prevent aggregation of the protein.

6.1.4.3. Purification of MED17-F1_His₆ (Srb4-F1_His₆)Table 17 – buffers for MED17-F1_His₆ purification

Buffer A	50 mM Tris-HCl pH 8.0, 10% (v/v) glycerol
Buffer B	1000 mM NaCl, 50 mM Tris-HCl pH 8.0, 10% (v/v) glycerol
Mono Q 5/5 HR	10-100% buffer B, 3 mM DTT
Superose12	30% buffer B, 3mM DTT

Cell pellets from 2-8 l expressions were lysed by sonication in 30 % buffer B with 10 mM β-mercaptoethanol and 1 mM Pefabloc (Table 17). Column washing was performed with lysis buffer containing 20 mM imidazole and elution with lysis buffer containing 300 mM imidazole. A Mono Q 5/5 HR anion exchange run was performed

followed by a gel filtration on Superose12. Protein solution was concentrated to 7.5 mg/ml for crystallization set-ups.

6.1.4.4. Purification of MED17-F2_His₆-bicistrons

Table 18 – buffers for MED17-F2_His₆ purification

Buffer A	50 mM Tris/HCl pH 8.0, 10% (v/v) glycerol
Buffer B	1000 mM NaCl, 50 mM Tris/HCl pH 8.0, 10% (v/v) glycerol
Wash	15% buffer B, 10 mM imidazole, 10 mM β-mercaptoethanol
Elution	15% buffer B, 200 mM imidazole, 10 mM β-mercaptoethanol

Cells from 500 ml liter expression culture were lysed in 15% buffer B containing 10 mM β-mercaptoethanol and 1 mM Pefabloc (Table 18). Cleared lysate was applied onto 500 μl Ni-NTA columns. After washing with lysis buffer a washing step with 10 mM imidazole was performed before elution.

6.1.4.5. Purification of the GST-MED17-F2/MED6ΔC_His₆

Table 19 – buffers for GST-MED17-F2/MED6ΔC_His₆ purification

Lysis buffer	150 mM NaCl, 50 mM Tris, 1 mM EDTA, 10% (v/v) glycerol, 2 mM DTT, 1x protease inhibitors mix
Buffer B	300 mM NaCl, 50 mM Tris, 20% (v/v) glycerol, 2 mM DTT, 1x protease inhibitors
Elution buffer	10 mM glutathione in buffer B
Ni column buffer	200 mM NaCl, 50 mM Tris, 10% (v/v) glycerol, 10 mM β-mercaptoethanol
Washes/Elution	0 mM, 10 mM, 20 mM and 300 mM imidazole in Ni column buffer

GSTMED17ΔN/MED6His₆ was purified by sequential Ni-NTA and GST affinity chromatography (Table 19). After sonication cleared lysate was applied on a glutathione-sepharose equilibrated with buffer B. After elution sample was diluted to Ni column buffer. Elution was performed by stepwise increasing the imidazole concentration, the complex started to elute at 20 mM imidazole.

6.1.4.6. Purification of MED20/MED18_His₆ (Srb2/Srb5_His₆) and MED8/MED20/MED18_His₆

Table 20 – buffers for MED20/MED18_His₆ and MED8/MED20/MED18_His₆ purification

Buffer A	20 mM Tris/HCl pH 8.0, 10% (v/v) glycerol
Buffer B	1000 mM NaCl, 20 mM Tris/HCl pH 8.0, 10% (v/v) glycerol
Mono Q 10/100 GL	10-100% (MED20/MED18_His ₆) and 10-50% buffer B (MED8/MED20/MED18_His ₆), 5 mM DTT
Superose12	15% buffer B, 5 mM DTT
Crystallization MED20/MED18_His ₆	2 M NH ₄ SO ₄ , 10 mM DTT
Crystallization MED20/MED18_His ₆	2 M NH ₄ SO ₄ , 100 mM NaAcetate pH4.6, 10 mM DTT
Crystallization MED20/MED18_His ₆	3.5 M NaFormate, 100 mM NaAcetate pH4.6, 10 mM DTT

Bacterial cells from 0.5-4 l expression were lysed by sonication 15% buffer A containing 10 mM β -mercaptoethanol and 1 mM Pefabloc and protein complexes were eluted from a Ni-NTA column with lysis buffer containing 300 mM imidazole (Table 20). After anion exchange purification and subsequent gel filtration complexes were concentrated to 4 and 8 mg/ml for crystallization trials, which in case of the dimeric MED20/MED18_His₆ (Srb2/Srb5_His₆) yielded microcrystals complex at the higher concentration.

The substoichiometrically present trimeric complex can be separated from the excess of dimeric complex by a very shallow gradient in the anion exchange protocol and as a separate peak eluting from the gel filtration.

6.1.4.7. Purification of the GST-MED17-F2/MED6ΔC MED8/MED20/MED18-His₆ coexpression

Table 21 – buffers for GST-MED17-F2/MED6ΔC MED8/MED20/MED18-His₆ purification

Buffer A	20 mM Hepes pH 7.5, 5 mM DTT
Buffer B	1000 mM NaCl, 20 mM Hepes pH 7.5, 5 mM DTT
Elution buffer	150 mM NaCl, 20 mM Tris/HCl pH 8.0, 10% (v/v) glycerol, 5 mM DTT, 50 mM glutathione pH8.0
Mono Q HR 5/5	100-1000 mM NaCl, 20 mM Tris/HCl pH8.0, 10% (v/v) glycerol, 5 mM DTT
Superose12	150 mM NaCl, 20 mM NaCl Tris/HCL pH8.0, 10% (v/v) glycerol, 5 mM DTT

Lysis by sonication of bacterial cells from a 4 l coexpression culture was conducted in 15% buffer B containing protease inhibitors. Elution from 1 ml glutathione sepharose was performed with a higher pH Tris buffer to increase elution efficiency (Table 21). After anion exchange and gelfiltration chromatography, samples were analyzed by SDS-PAGE and resulting band cleaved and verified by mass spectrometry.

6.1.4.8. Purification of CDK8_His₆

Table 22 – buffers for CDK8_His₆ purification

Buffer A	137mM NaCl; 2,7mM KCl; 4,3 mM Na ₂ HPO ₄ •7H ₂ O; 1,4mMKH ₂ PO ₄ ; pH 7.4, 1x protease inhibitor mix, 1mM NaF, 0.4 mM Na ₃ VO ₄ , 3 mM DTT
Buffer B	like buffer A with 1000 mM NaCl
Buffer C	like buffer A with 500 mM imidazole pH 7.4

Sf9 cells from 100-500 ml infected cultures were harvested by centrifugation (10 min, 300 g, 4°C) resuspended in buffer A, and lysed by sonication for 40 sec (Table 22). The lysate was cleared by 1h at 100.000 g in a swing out bucket rotor (4°C) and the cleared lysate was loaded onto a Ni HiTrap column, which was then

connected to the Äkta purifier system in order to perform two subsequent gradients: 0-100% buffer B and 0-100% buffer C. CDK8_His₆ is eluted during the imidazole gradient. For best results it is advisable to run a slow imidazole elution gradient. A Mono S cation exchange was performed with 0-100% buffer B.

6.1.4.9. Purification of Cyclin C

Table 23 – buffers for Cyclin C purification

Buffer A	20 mM Tris/HCl pH 8.0, 10% (v/v) glycerol
Buffer B	1000 mM NaCl, 20 mM Tris/HCl pH 8.0, 10% (v/v) glycerol
Mono Q	10-60% buffer B, 5 mM DTT
Superose (6 or 12)	10% buffer B, 5 mM DTT
Crystallization	200 mM magnesium formate, 10 mM DTT
Crystallization	1250 mM NaAcetate, 100 mM imidazole pH 6.0, 7.5% PEG 400, 10 mM DTT

Cells from 2 l expression culture were lysed in 30% buffer B containing 10 mM β -mercaptoethanol and 1 mM Pefabloc (Roth). Cleared lysate was applied to 1.5 ml Ni-NTA resin (Table 23). After the elution profile was determined with an imidazole gradient in the lysis buffer, washing routinely included 2 steps with 2ml 10 mM and 2 steps with 2ml 70 mM imidazole in the lysis buffer. Elution fractions were diluted with buffer A containing 10 mM β -mercaptoethanol and subjected to Mono Q 5/5 HR ranging from 10% buffer B to 60% buffer B containing 5 mM DTT over 10 CV, followed by concentration to 1000 μ l total volume and two separate gel filtration steps on a Superose column with 10% buffer B containing 5 mM DTT. For crystallization Cyclin C was concentrated to approx. 4 mg/ml.

6.2. Limited Proteolysis

For chymotrypsin and trypsin treatment 0.25 μ g of protease per gel sample was added to the total protein mixture in the reaction buffer. At each time point a sample of 50 μ l was taken out and pipetted into a prepared tube with 10 μ l 6x sample buffer to be boiled immediately. For subtilisin or proteinase K treatment protein

samples were pipetted into separate tubes and 1 μ l of protease mixture at various concentration (430 ng/ μ l, 43 ng/ μ l, 4.3 ng/ μ l and 0.43 ng/ μ l for proteinase K and 1 μ g/ μ l, 100 ng/ μ l, 10 ng/ μ l and 1 ng/ μ l for subtilisin, empirical values) was added. The mixture was incubated for 1 h on ice before adding 10 μ l 6x sample buffer and boiling. Ideally the protein concentration was adjusted to yield around 50 μ g of protein per sample. In case these amounts were not at hand, silver staining of gels was conducted to determine a time/concentration point where intermediary products are visible and perform the experiment in a second round only with the determined incubation time. In this case the samples were precipitated with -20 °C acetone and pellets resuspended in 1x sample buffer before SDS-PAGE. Bands visible in a Coomassie-stained gel were either reproduced and blotted onto PVDF membrane to be cleaved or cleaved directly and transferred to PVDF membrane by the passive transfer protocol before Edman sequencing with Procise® Protein Sequencing System (Applied Biosystems).

6.3. Determination of protein-protein interactions

6.3.1. GST Sepharose® pull-down assay

For GST pull-down assays 1.5 ml cleared lysates from GST-CTD expression and GST alone control expression were incubated for 4 h on a turning wheel at 4°C with 30 μ l of equilibrated GST Sepharose® beads (Amersham). Four washing steps, consisting in 1ml washes twice with lysis buffer, once lysis buffer with 1000 mM NaCl and once with lysis buffer followed by a 1 min centrifugation at 300g, 4°C were performed. After washing the bound protein was estimated with a Bradford assay and the amount of resin solution was normalized to equal concentrations of GST. A resin alone control was prepared for the second incubation step. In this step 60 μ g of purified Cyclin C was added to the beads in 1.5 ml volume of lysis buffer containing biotin (300 μ g) and varying amounts of NP40. The incubation was performed for four hours on the turning wheel at 4°C. Three washes were performed with the corresponding buffer conditions without Cyclin C before 40 μ l of 1x sample buffer was added, samples were boiled and analyzed by SDS-Page and Western-blotting.

6.3.2. Microcalorimetry

Temperature changes upon consecutive injections of 15 μ l peptide solution (50 μ m) into 1.4 ml of a 5 mM Cyclin C solution (in 100 mM NaCl, 20 mM Tris pH8.0, 10% glycerol, 5 mM β -mercaptoethanol) measured in a isothermal microcalorimeter (VP-ITC, MicroCal) For accuracy protein and peptide concentrations were calculated from the absorption at 280 nm. As a control temperature changes upon injection of buffer only into the Cyclin C solution was measured.

7. SF9 insect cell culture and recombinant baculovirus expression techniques

7.1. Growth conditions for SF9 cells

Insect cells were grown in SF900 II medium without antibiotics at 27°C when not infected and at 28°C when infected with virus. Doubling time of healthy cultures was usually approx. 24h and was determined with a Neubauer improved counting chamber (Peske). Unless for subsequent freezing or in case of slower doubling times or apparent system malfunctions cells were grown without FCS, for bigger expression penicillin/streptomycin was added. Cells were split 1:3 into fresh flasks approx every 2-3 days when culture plate or flask became 100% confluent.

7.2. Freezing and thawing of insect cells

For long term storage and strain establishment, insect cells were harvested at 80% confluence in mid-log phase, centrifuged at 300 g for 10min and resuspended in medium containing 10% DMSO and aliquoted. Subsequently cells were slowly cooled down for 1 h at -20°C before transferring them to -80°C and finally into liquid nitrogen. For thawing, a 1 ml liquid nitrogen stock was quickly submerged in a 37°C water bath until it was completely thawed (1-2 min) avoiding shaking. The culture was quickly pipetted into tissue culture 25 ml flask with 5 ml preheated medium using a sterile plastic Pasteur pipette to avoid shearing forces that would occur with pipette tips with smaller outlets. Cells were allowed to settle during 1 h at room temperature with the colonizable surface of the flask being reduced by creating an inclined plane. After 1 h medium was carefully removed and fresh preheated medium was provided, cells were transferred into the 27°C incubator. Subsequently medium was exchanged

several times on the same day and then once a day until the plate was confluent again and split with normal rates.

7.3. Transposon mutagenesis and blue/white selection

The pFastBac1-CDK8 plasmid was transformed into DH10-BacTM cells. Instead of directly plating the cells after heat shock recovery (in 1 ml SOC medium), 10 ml of SOC medium supplemented with kan, tet and gen were added and cells were grown for 96 h under vigorous shaking (230 rpm) at 37°C before plating them at three different concentrations (1:1, 1:10 and 1:100) onto fresh Blue/White selection plates containing the antibiotics plus IPTG and X-Gal. Plates were incubated overnight at 37 °C. Suitable white colonies were picked and replated on selection plates for a color control. A glycerol stock was prepared (addition of 20% sterile glycerol in 1 ml total volume, storage at -80°C) and the DNA was isolated as described above with the user-adapted protocol for MidiPrep with the Qiagen Kit. Success of mutagenesis was controlled by PCR using combinations of M13 primers, CDK8 specific primers, and combinations of both.

7.4. Isolation of Bacmid DNA

For the isolation of Bacmid DNA, a 100 ml culture with selective LB medium was inoculated with 0.5 ml from a 5 ml overnight starter culture of recombinant DH10 Bac cells and grown for 14 h under vigorous shaking (250 rpm). DNA was isolated with the DNA MidiPrep Kit according to a user adapted protocol where the QF elution buffer was preheated to 65°C and elution was performed in five 1ml steps instead of one 5ml step. This leads to a higher yields of high molecular weight DNA, the division in 5 steps prevents cooling of the buffer. Bacmid DNA was resuspended in 100 µl TE buffer and had a concentration of 580 µg/ml.

7.5. Transfection of SF9 cells with Bacmid DNA

15 µl of Bacmid DNA at 581 ng/µl and 18 µl of Lipofectin (Invitrogen) were each mixed thoroughly with 250 µl of serum free medium without antibiotics in sterile tubes made of polystyrene (Falcon). Usage of a different material greatly reduces transfection efficiency as DNA-lipid complexes stick to the reaction tube when this is made for example of polyethylene. The two solutions were combined and incubated

at room temperature for 15-45 min. For transfection 2 ml of medium was added to the transfection mix, an 80% confluent mid-log phase 5 ml insect culture was washed once with medium and overlaid with the mix. Cells were incubated at 28°C for 5 hours before the transfection mix was aspirated off and overlaid with 5 ml medium containing antibiotics. Cells were incubated for 5 days and supernatant was harvested by centrifugation (10 min, 300g, 4°C).

7.6. Harvesting of initial virus stocks and virus reamplification

The 5 ml first generation virus stock was split in two and two 80% confluent mid-log phase 5 ml insect cell cultures were reinfected overlaying the culture with 4 ml fresh medium and adding 2.5 ml virus stock to each flask. Cells were incubated until lysis occurred (5 days) and supernatants were harvested by centrifugation. In a third and fourth step culture sizes were increased by diluting the virus 1:4 and 1:10 respectively and infection were carried out for 48 h. Supernatants from subsequent expression tests (48-72 h) were pooled and used as the working virus stock. An optimal infection ratio of 1:7 where 90% of the cells had detached 24 h postinfection was determined for this stock. Supernatants of subsequent infections were pooled for a new stock. Virus stocks were stored at 4°, where they are stable for a few weeks.

7.7. Expression of CDK8

Expression was carried out applying fresh medium mixed at a 1:7 ratio with infectious supernatant onto an 80% confluent plate that had been split the previous day. The optimal infection time was determined and further expression were carried out for 24 h before cells were harvested by centrifugation at 4°C and resuspended in lysis buffer. Supernatant was removed and reused as described.

8. Electrophoretic methods

8.1. Electrophoretic separation of DNA

DNA was separated in horizontal TBE 1% agarose gels; the ethidium bromide concentration in the gels was 0.7 µg/ml. Samples were mixed with 1/10 vol. of sample buffer. For visualization of DNA the gel was placed on a UV-screen (Eagle Eye™, Stratagene)

8.2. Protein separation by SDS-PAGE

For proteins samples glycine-SDS-PAGE with 10%-17% acrylamide gels (Laemmli, 1970) was performed. Gels were then either subjected to protein transfer for Western-blot or directly stained with Coomassie solution. Silver staining was performed (1h fixing, 2x10min washing, 1min sensitizing, 3x30sec H₂O wash, 20sec staining, 2x 30sec H₂O wash, development, reaction stop) when very low amounts of proteins had to be visualized.

9. Immunological methods

9.1. Protein transfer and Western blot

After SDS-Page proteins were blotted in a transfer chamber (BioRad Trans-Blot Cell) for either 1 h at 90 V, 3 h at 60 V or 16 h at 20 V onto PVDF membranes (Roth), prewet with ethanol. Blotting steps included 1 h blocking of unspecific interactions and 1 h incubation time for each antibody. Washing steps were carried out after each antibody incubation for 15 min followed by two times 10 min. Either the secondary or primary antibody was coupled to HRP (horse-radish-peroxidase) and detection was performed with the ECL Plus reagents (Amersham Pharmacia Biotech) following the manufacturers instructions.

9.2. Passive adsorption method for protein transfer

Alternatively proteins can be transferred to PVDF membrane by passive adsorption. This is particularly useful for Edman sequencing because it avoid N-terminal end blocking produced by electric transfer. Coomassie stained bands are directly cut from the gel and dried in a Speed Vac. Dried pieces were swollen in 50 μ l swelling buffer. After addition of 200 μ l H₂O a 1-5 mm² piece of prewet (methanol) membrane was added. Methanol was added to 10% final concentration after the solution had began to turn blue. After 1-2 days, membrane was washed by vortexing five times with 1 ml 10% MeOH for 30 sec.

10. Crystallization

Crystal setups were performed either manually with the hanging drop vapor diffusion method or with the crystallization robot (Hydra Plus 1) as sitting drops in 96-well plates at 20°C. Full-length Cyclin C was crystallized with the hanging drop method using as reservoir solution 250 mM magnesium formate, 25% glycerol, and 10 mM DTT, reached a maximum size of 70 x 70 x 70 μm , and could be flash-cooled directly from the hanging drop. The variant comprising residues 5-228 was highly soluble and could be crystallized using a reservoir solution containing 1.25 M sodium acetate, 100 mM imidazole pH 6.5, 7.5% PEG 400, and 10 mM DTT with a protein to crystallization solution ratio of 2:1. Crystals were harvested in the reservoir solution. The PEG 400 concentration was increased to 25% in six steps using an incubation time of one hour after each step. Crystals were slowly cooled to 4° in a styrofoam box, incubated overnight, mounted into nylon loops (Hampton research), and were plunged into liquid nitrogen for storage.

11. X-Ray analysis

11.1. Data collection

Data were collected at the protein crystallography beam line X06SA at the Swiss Light Source (SLS), Villigen/Aargau. This was necessary since the crystals showed a rather low diffraction power. Additionally they were very radiation sensitive.

Therefore a rotation of 0.5° per frame was chosen, which would allow obtaining good $I/\sigma(I)$ values with a low x-ray dose. This strategy was combined with a beam focused on the detector rather than on the sample. Radiation damage was observed in a reduction of high angle intensities due to disorder. Since radiation damage usually continues at a steady state rate, even if exposure ceases (probably due to effects initiated by free radical), intensity measurements were carried out rapidly and without interruption. Since the signal-to-noise ratio increases with the distance, the maximum reasonable crystal to detector distance was chosen for data collection.

In order to achieve maximum redundancy (total number of reflections measured/total number of unique reflections) and a high completeness (total number of intensity measurements/total number of measured unique reflections), the total

oscillation range of 100° as well as the starting point was determined prior to data collection with MOSFLM (Powell, 1999).

Multi wavelength anomalous dispersion (MAD) datasets were recorded from selenomethionine-containing crystals. In order to exploit the changes in both the normal scattering (f_H) and the anomalous scattering (f_H') of the anomalous scatterer (selenium) three datasets with 100° each were measured at wavelengths close to the absorption band. Peak data at a wavelength of 0.97977 \AA were recorded first, then data at the inflection point at 0.98004 \AA and remote data at 0.94927 \AA .

12. Data processing and structure solution

12.1. Indexing, integration and scaling

Data were processed with DENZO and SCALEPACK (Otwinowski and Minor, 1996), except the native dataset, which was processed with MOSFLM and SCALA (CCP4, 1994) and data from peptide soaking experiments, which were processed with HKL2000 (Otwinowski and Minor, 1996). For the MAD data the error model was adjusted for the remote data set and applied to the two remaining datasets.

12.2. Phasing

Program SOLVE (Terwilliger, 2002) was used for MAD phasing with the datasets from selenomethionine containing crystals. Phases were improved with SHARP (La Fortelle and Bricogne, 1997). The resulting electron density map was used for positioning a polyalanine model of the cores of the two helical cyclin repeats. Subsequent phase combination and phase extension to 3.0 \AA with amplitudes from native data allowed building of an atomic model. For cocrystallization and peptide soaking experiments, phases were obtained by molecular replacement with the Cyclin C structure using the program PHASER (Storoni *et al.*, 2004) followed by a rigid body refinement. In order to avoid loss of a weak additional density bulk solvent correction was disabled in this and all following refinement steps.

12.3. Refinement

Refinement of the structure with CNS (Brunger *et al.*, 1998) consisting of rounds of application of the routines MINIMIZE and BINDIVIDUAL. Quality of the

resulting model was analyzed with 2Fo-Fc and Fo-Fc electron density maps. Stereochemistry was judged with PROCHECK (Laskowski *et al.*, 1993b).

Results

1. Mapping of subunit domains and subunit-subunit interactions within the Mediator head module

The Mediator head module, which consists of eight subunits, tightly binds to RNA Pol II (Davis *et al.*, 2002). Determining a detailed interaction map of stable subcomplexes within the Mediator head module is the first step to achieve high expression levels of recombinant subcomplexes. 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 multiprotein context. To overcome this obstacle and to analyze at the same time subunit interactions within the Mediator head module, a coexpression strategy was combined with limited proteolysis studies. Domain mapping through limited proteolysis and bioinformatical tools led to truncated and more stable variants of the subunits. Subsequently stable variants were coexpressed and copurified with binding partners. Iterative proteolysis and truncation of the coexpressed and copurified subunits allows the determination of stable subcomplexes. This approach may be used to obtain potentially crystallizable portions of other multiprotein complexes.

1.1. MED17 (Srb4), the integral subunit of the Mediator head module

The MED17 (Srb4) protein binds several subunits within the Mediator head module (ref). This 75 kD protein is therefore regarded as a scaffold protein in the head module, making it as much an interesting as difficult target for structural studies. The Srb4 structure would provide information on the organization and function of this module. Srb4 is very difficult to express alone.

As a first attempt to obtain MED17 in soluble form overexpressed GST-MED17 (Srb4) fusion protein was purified in a three-step protocol. It consisted in an affinity chromatography over glutathione-sepharose, a MonoQ anion exchange chromatography and a final gel filtration step over a Superose12 column. The yield of protein did not exceed 50 µg from a 4 liter culture. From gel filtration it was clear that most of the protein was aggregated since it all eluted in the void volume. The expression protocol for MED17 (Srb4) could still be improved by optimizing time,

temperature and expression strand. The resulting standard protocol consisted of expression over night at 18°C with the BL21RIL strand. In order to prevent the protein from aggregating, experiments were carried out in different buffer and salt conditions. Despite some improvement, purification would be very limited by low yields and copurification of the chaperone DnaK (confirmed with Edman sequencing), a fact underlining the partially misfolded nature of the recombinantly expressed protein.

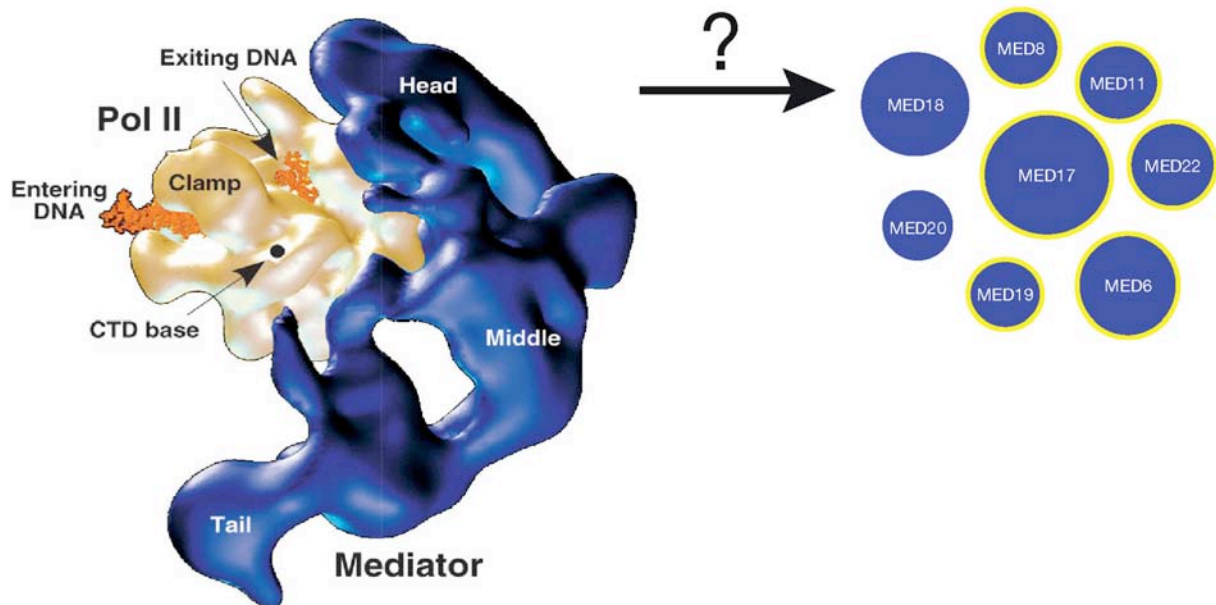


Fig. 2 – Overall structure of RNA Pol II-Mediator complex and subunits of the head module. The Mediator of transcriptional activation binds extensively to RNA Polymerase II as it can be derived from EM- analysis (left picture, Davis *et al.*, 2002). The head module is thought to be the major RNA Pol II interacting module. It is not clear where and by which of the eight subunits this binding occurs, nor has it been proven how they are organized among each other. A yellow circle lines the subunits essential in *S. cerevisiae*.

Limited proteolysis revealed two stable domains of MED17 (Srb4), referred to as F1 and F2 fragment. The complete MED17 (Srb4) fraction from a large-scale expression was subjected to partial proteolysis using chymotrypsin and trypsin, to probe for flexible regions that may interfere with solubility and stable domains suitable for a high yield expression. Two cleaved fragments starting with amino acid 97 and 241 were identified by Edman sequencing (Fig. 3A). This information, combined with secondary structure prediction, allowed the design of several new constructs and variants, which were tested for their expression levels and solubility. The resulting MED17 (Srb4) domain structure, the secondary structure prediction,

conserved essential regions, as well as important sites determined by other groups are summarized as a map in figure 3B.

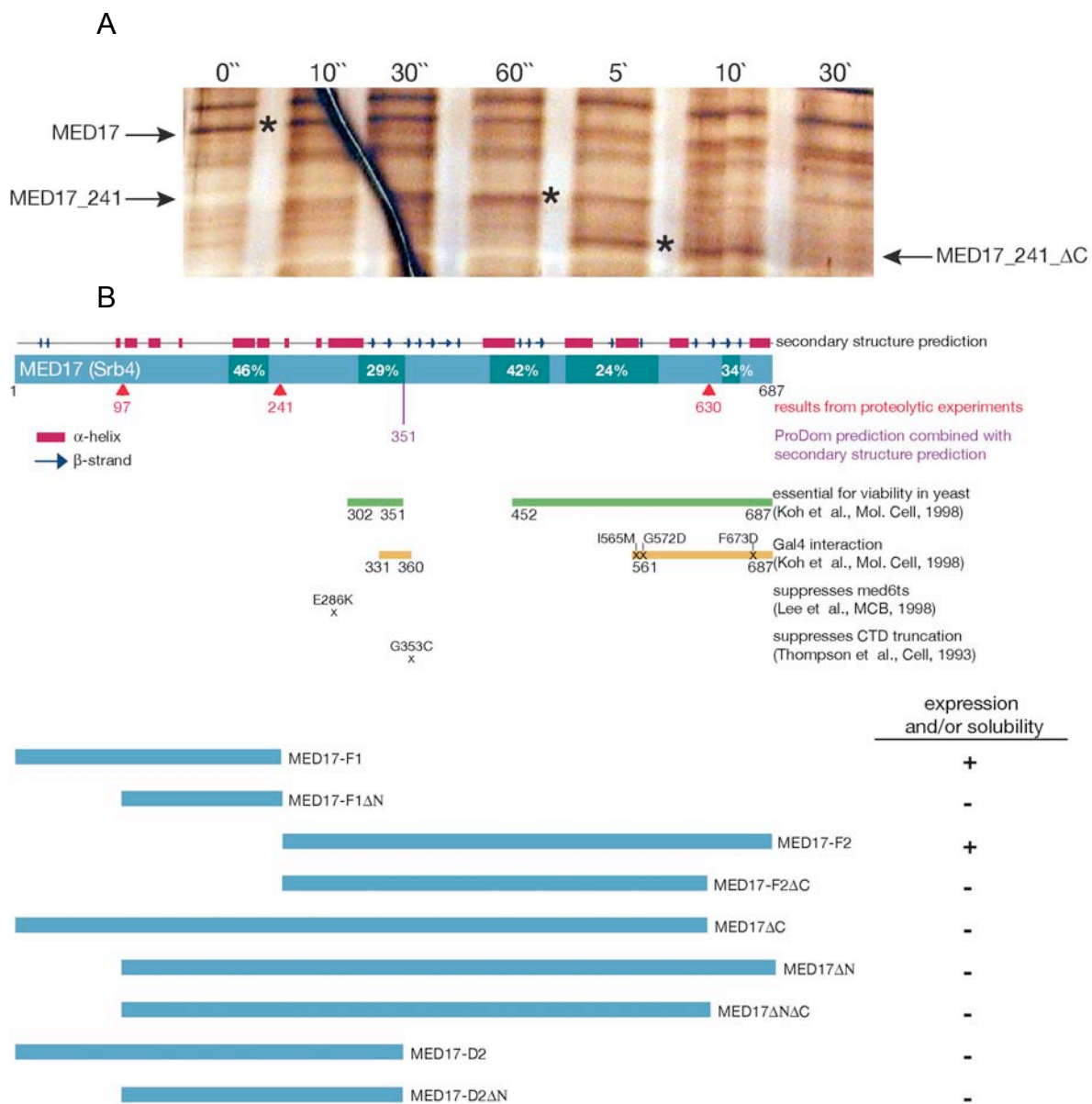


Fig. 3 – Analysis of MED17 (Srb4) domain structure. (A) Limited proteolysis using chymotrypsin reveals a cleavage site at amino acid 241 and an estimated site at the C-terminus. Full-length MED17 (Srb4) and the two intermediary products MED17-F2 (residues 241-687) and MED17-F2ΔC (residues 241-630) that appear 30'' and 5' respectively after addition of protease are marked with a star. (B) Summary of the results on MED17 (Srb4) domain analysis. Results from previous genetic and biochemical studies are listed along with results from secondary structure prediction and limited proteolysis. Two constructs designed according to these results led to soluble protein variants. The other protein variants displayed, if at all, very low expression levels, indistinguishable from background in affinity purifications.

1.2. Two fragments of MED17 (Srb4) display better expression and solubility than full-length protein

Most of the variants of MED17 (Srb4) could not be overexpressed (Fig. 3B). Only the two major products of the proteolysis assay led to soluble variants, MED17-F1 and MED17-F2. However, MED17-F2 was poorly expressed and difficult to purify. To prevent aggregation, a protocol using high salt concentrations for the gel filtration chromatography had to be established, which led to some improvement. Neither the GST-fused MED17-F2 nor a version with a C-terminal hexahistidine tag could be purified in amounts suitable for crystallization. In contrast, the short MED17-F1 with a C-terminal hexahistidine-tag was expressed in soluble form in reasonable amounts, a purification protocol was established, and a CD spectrum indicated proper folding (data not shown). Crystallization screens using standard *Hampton Research* conditions did not yield crystals.

1.3. Bicistronic expressions of MED17 (Srb4) with the other Mediator head module subunits reveals binding to MED6 and MED8

The main obstacles for reconstitution of Mediator head module are poor expression levels and insolubility of *E. coli* expressed proteins. The loss of structural integrity from which individual subunits suffer outside their natural multiprotein context can be overcome by multicistronic expression strategies. A multicistronic mRNA contains individual ribosomal binding sites (RBS) for each gene. When the mRNA is translated the nascent polypeptide chains of partner proteins are synthesized in close proximity, favoring immediate formation of the stabilizing complex. Subsequently coexpressed subunits are copurified.

Thus data obtained from proteolysis and domain mapping can be used in attempts to improve expression and solubility by coexpression with potential binding partners. At the same time such copurification can successfully map strong and specific direct protein-protein interactions, as demonstrated by subsequent structure solutions of such protein complexes (Baumli *et al.*, 2005). Specific interaction is indicated by the fact that only one protein carries an affinity tag and the potential interaction partner can be copurified. 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 persist over several copurification steps.

1.3.1. Screening with bicistronic vectors reveals binding of MED17 to MED8

For screening protein-protein interactions in the Mediator head module bicistronic expression vectors were designed systematically by combining the MED17-F2 fragment with other subunits of the Mediator head module. Except MED6 and MED8 none of the subunits could be copurified with MED17-F2 (Fig. 4).

As expected for a scaffold protein, offering just a single expression partner did not fully overcome problems in expression and solubility. No effect could be seen in combination with MED19 (Rox3), MED20 (Srb2), MED11, and MED22 (Srb6) (not shown). Apart from not binding at all to MED17 (SRB4), some of them might also bind the conserved region in the MED17-F1 fragment (Fig. 3B). Determination of a recombinantly expressed and soluble subcomplex out of this dense network of interactions will need further combinatorial trials to define the minimally required interactions.

In this screen MED8 was determined for the first time an interaction partner of the MED17-F2 fragment. MED17-F2 yields are much higher and an additional stoichiometric binding partner is detected in the Coomassie stained gel when compared with the other co-purification assays (Fig. 4A). The protein band was excised from the gel and identified by mass spectrometry as MED8. MED8 is thus stoichiometrically copurified in an affinity purification targeting only MED17 (SRB4). Copurification was confirmed at a later stage in large-scale expressions, but the tendency of MED17-F2 and maybe also MED8 to aggregate prevented high yield purification results.

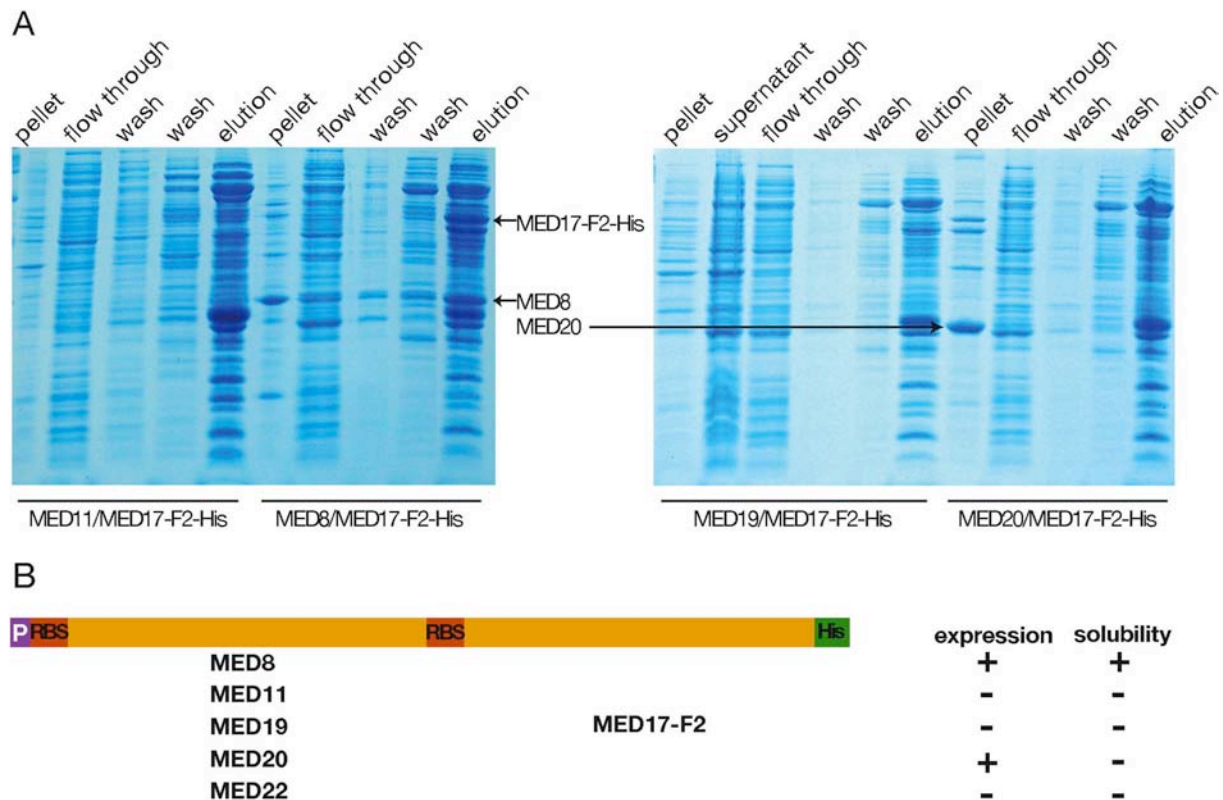


Fig. 4 – Screen of bicistronic expression within the Mediator head module. (A) Affinity purification of coexpression tests of MED17-F2-His₆ with MED8, MED11, MED19, MED20. Bands that could be identified (by Western Blot, Edman sequencing, or mass spectrometry) are marked with an arrow. (B) Summary on results of bicistronic expression with MED17-F2-His₆. Expression and solubility in the individual expression tests are indicated.

1.3.2. MED6 bridges the two Mediator core modules

It was shown (Baumli *et al.*, 2005) that the MED7/MED21 heterodimer binds directly to MED6, which shows 34% sequence homology between yeast and human (Fig. 5A). MED6 is an integral part of the head module (Lee and Kim, 1998), suggesting that MED6 bridges between these two modules. To test if MED6 binds directly to MED17 (SRB4), the architectural subunit of the head module (Koh *et al.*, 1998), MED6 was tagged with a C-terminal hexahistidine tag (His₆), fused MED17 (SRB4) to a N-terminal GST tag (GST), and the two subunits were coexpressed from a bicistronic vector. In two subsequent affinity chromatography steps, using a Ni-NTA and a glutathione column, the complex was copurified (Fig. 5A). Successful purification of the complex was independent of the order of the affinity columns (Fig. 5B). The weakly conserved N-terminal part of MED17 (SRB4), MED17-F1, and the non-conserved C-terminal part of MED6, are not required for binding since truncated variants of MED17 (MED17-F2, residues 241-688) and MED6 (residues 1-214) were

sufficient for the interaction. These results are consistent with a functional interaction between MED17 (SRB4) 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 (SRB4) in the head module and with the MED7/MED21 heterodimer in the middle module.

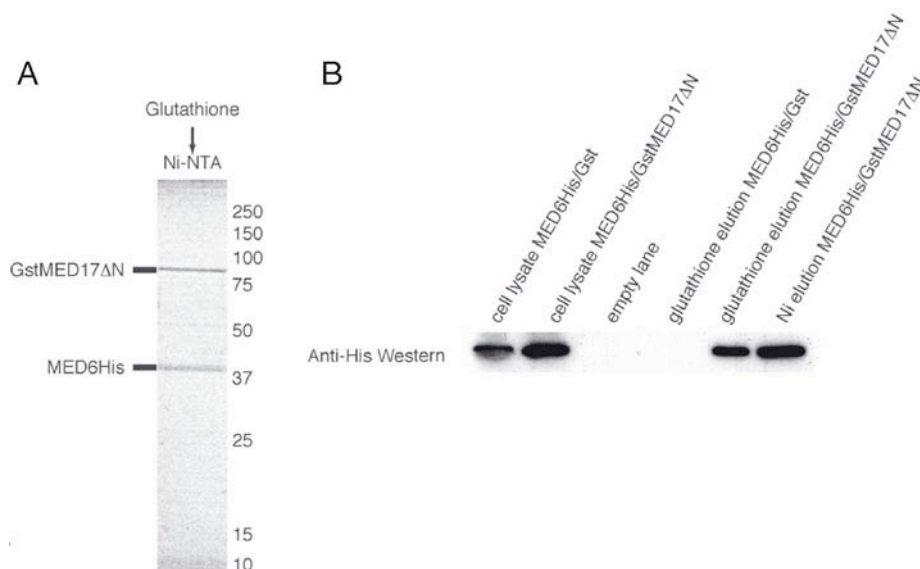


Fig. 5 - Copurification of MED6 with MED17 (Srb4). (A) copurification of MED6-His₆ with GST-MED17 Δ N. A schematic presentation of the purification procedure is shown above the Coomassie stained gel. (B) Western blot of the GST-MED17 Δ N/MED6-His₆ purification. Binding of MED6 to MED17 is not due to the presence of the GST tag since a purification using GST only does not yield MED6 (fourth lane).

1.4. A trimeric subcomplex MED20/MED18/MED8 in the Mediator head module

1.4.1. Preparation of soluble MED20/MED18 (Srb2/Srb5) - heterodimer

The MED20 (Srb2) and MED18 (Srb5) proteins are non-essential in yeast but highly conserved subunits of the Mediator head module. In contrast to expression of MED20 (Srb2) in a bicistronic vector with MED17 (Srb4) (Fig. 4), which results in insoluble protein, offering MED18 (Srb5) as a binding partner yields high amounts of soluble protein, demonstrating the effectiveness of this method as soon as the right interaction partners are brought together (Fig. 6A, upper construct). High yields (10 mg pure protein from 500 ml of *E. coli* culture) of the complex were copurified stoichiometrically over a Ni-NTA affinity column using a C-terminal His tag only on MED18 (Srb5), a MonoQ anion exchange column and a Superose gelfiltration

column (Fig. 6B, left panel). The complex was stable in solution at 4°C for several weeks.

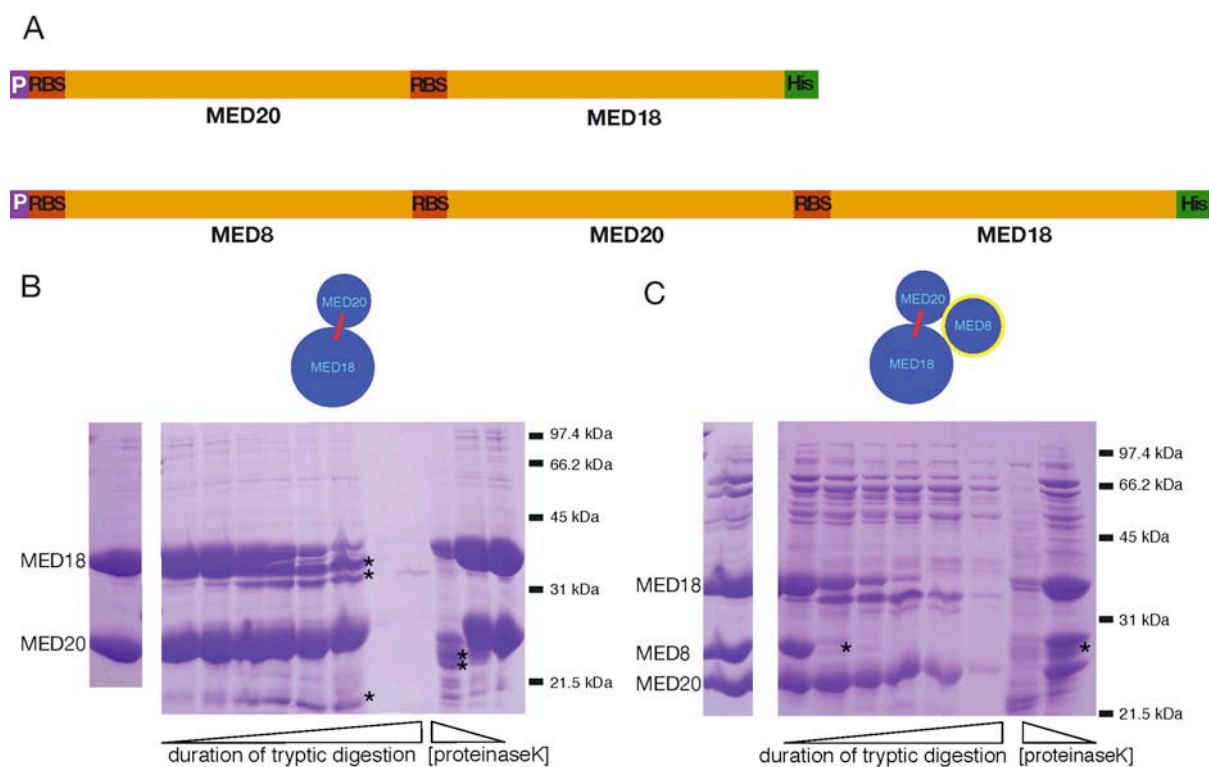


Fig. 6 – Stable subcomplexes within the Mediator head module: MED20/MED18 form a stable dimer that interacts with MED8. (A) multicistronic expression constructs. (B) purification and limited proteolysis of the dimeric MED20/MED18 complex and (C) the trimeric MED20/MED18/MED8 complex. The left panel of (A) and (B) shows the complex as it was purified in a three-step protocol. Tryptic digestion and proteinase K digestion were performed with both complexes. Stable intermediary products that were subsequently analyzed by N-terminal sequencing (see Fig. 8A and text) are marked with a star.

1.4.2. Crystallization of MED20/MED18 (Srb2/Srb5) heterodimer

Crystallization screens were performed at two concentration of the MED20/MED18 complex using standard *Hampton Research* screens. Microcrystals grew in two conditions (2 M NH_4SO_4 , 10 mM DTT and 2 M NH_4SO_4 , 100 mM NaAcetate pH4.6, 10 mM DTT), but fine screens did not yield any improvement (Fig. 7). Additional screens performed at higher salt concentrations yielded additional microcrystals, of which those grown in 100 mM NaAcetate pH4.6, 3.5 M Na formate, 10 mM DTT looked worth further refinement. The majority of the drops remained clear even at very high salt concentration and a protein concentration of 40 mg/ml.

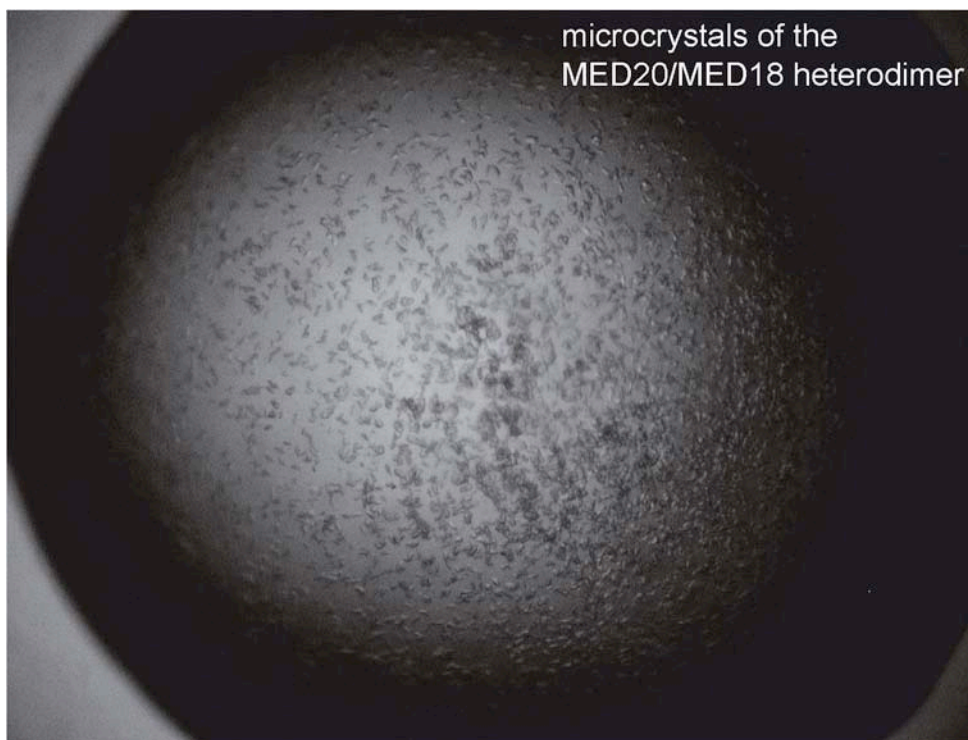


Fig. 7 – The MED20/MED18 heterodimer forms microcrystals at 8mg/ml in 2 M NH_4SO_4 , 100 mM NaAcetate pH4.6, 10 mM DTT. Hanging drop (1.5 μl : 1.5 μl) crystallization was performed with the dimeric complex at two protein concentrations. A picture was taken from a drop containing microcrystals of around 5 μm lengths.

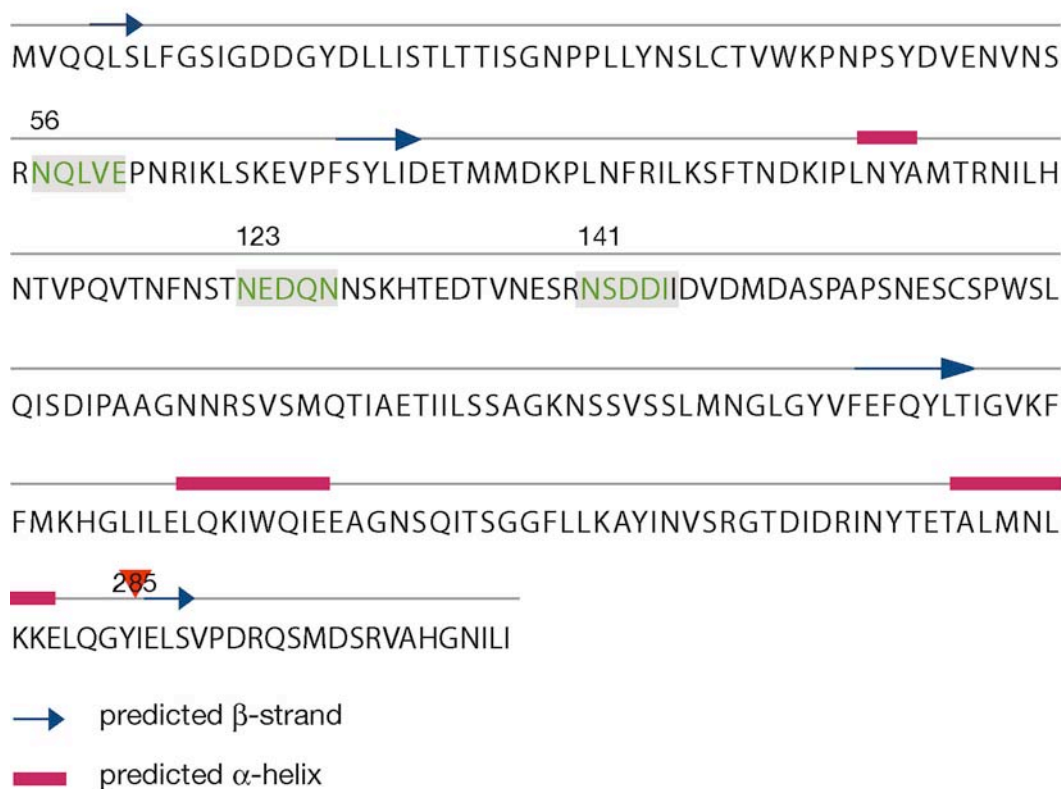
1.4.3. MED8 binds to the MED20/MED18 (Srb2/Srb5) heterodimer

In a first attempt to construct a tricistronic vector the *med8* gene was introduced with an additional RBS into the bicistronic vector containing the MED20 (Srb2) and MED18 (Srb5) genes (Fig. 6A, lower construct). After affinity chromatography and anion exchange chromatography, MED8 elutes with MED20/MED18 (Srb2/Srb5) in a stoichiometric trimeric complex from the gelfiltration column (Fig. 6C, left panel), proving the interaction. A big excess of the dimeric complex elutes from the gelfiltration column before the trimeric complex (not shown). This indicates a low stability of the MED8 protein. No crystals could be obtained from the trimeric complex in standard screens. In conclusion MED8 forms a complex with MED20/MED18 (Srb2/Srb5).

1.4.4. Limited proteolysis reveals potential variants of MED18 (Srb5) and confirms the high sensitivity and instability of the MED8 protein

In order to screen for stable domains in the dimeric as well as in the trimeric complex, both complexes were subjected to limited proteolysis experiments with subtilisin (Fig. 6B,C right panels) and trypsin (not shown).

A



B

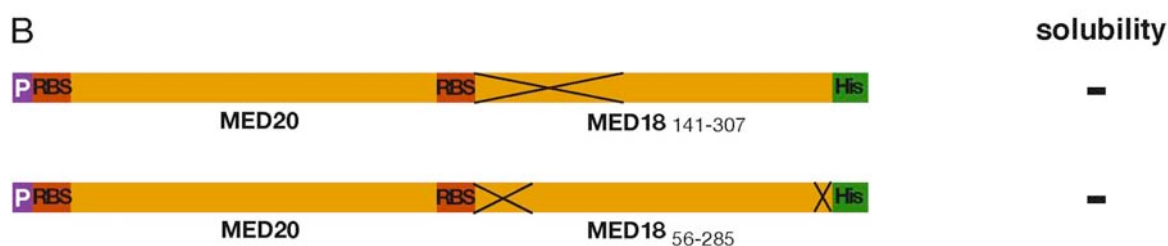


Fig. 8 – MED18 contains several protease sensitive sites. (A) Results from Edman sequencing of protease cleavage products are highlighted in the MED18 amino acid sequence. A number identifies their sequence position. An estimated C-terminal cleavage site is indicated by a triangle. The predicted secondary structure is outlined above the sequence with helices represented as rectangles and β -strands as arrows. (B) New bicistronic constructs that were designed based on limited proteolysis results led to insoluble proteins.

Consistent with its behavior in the purification of the stoichiometric trimeric complex, MED8 is a very unstable subunit in this subcomplex context. It is degraded at very low protease concentration and needs further binding partners for protection, possibly MED17 (Srb4).

Three N-terminal cleavage sites on MED18, amino acids position 56, 123, and 141, were identified with limited proteolysis and subsequent N-terminal sequencing of the resulting stable fragments. A C-terminal cleavage site, amino acid position 285, was estimated combining the size of the stable fragments with secondary structure position (Fig. 8A). Cleavage results for MED18 (Srb5) matched well with a secondary structure prediction (Fig. 8A). Variants of MED18 were subcloned in bicistronic vectors with MED20 (Srb2) and tested for behavior. All of them turned out to be insoluble (Fig. 8B), indicating either that cleavage sites consist in internal loop regions or binding with the partner molecule MED20 (Srb2) was impaired by truncation.

1.5. Design of a new tricistronic vector

The first generation of tricistronic vector was not very useful for further screening of interactions in the Mediator head module or other multiprotein complexes. The fastest way to test the tricistronic expression principle had involved duplication of a multiple cloning site rendering this construct non suitable for further gene exchanges. In order to easily exchange genes in this arrangement a vector was needed that allows easy switching between bi- and tricistronic expression and offers a variety of cloning sites for each of the three genes to provide higher flexibility. As in the previous bi- and tricistronic systems the last genes would be fused to a C-terminal His₆-tag. The complete cassette of the second and the third gene can be excised with *Nde*I /*Not*I and thus be introduced into a new vector in order to give rise to a bicistronic construct and vice versa. Since the restriction sites are compatible with the GST-fusion vectors constructed in the lab, a switch into a system with a different tag would also be easy to realize. Figure 9 provides a generalized map of the new vector. It was constructed in several variants with alternative restriction sites. Detailed information on sequences, restriction sites, and resistance cassettes is provided in Materials and Methods (table 5, tricistron II) and as supplementary material.

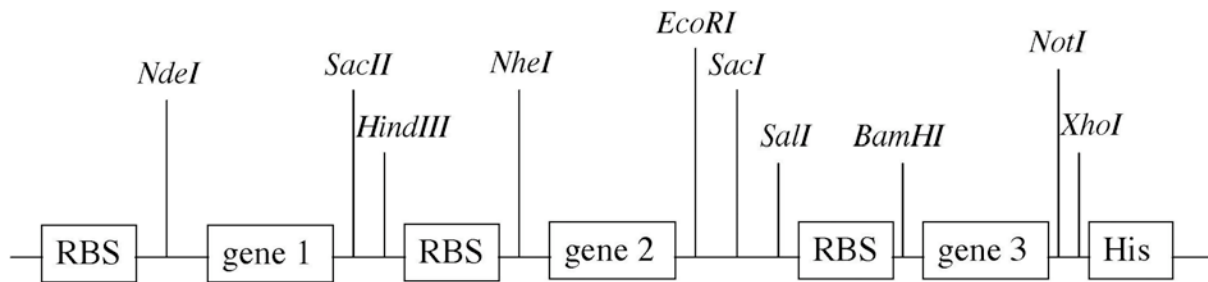


Fig. 9 – Outline of the vector construct for tricistronic expression. The vector, based on a pET24b vector (Novagen), was constructed by ligating PCR products that contained additional restriction sites, ribosomal binding sites and genes of interest. An alternative variant contains an *NcoI* site instead of the *BamHI* site. By cleaving with *NheI* and *NotI* a bicistronic construct can be transferred directly from or to the bicistronic vectors.

1.6. MED18 binds to MED8

The first genes cloned into the new tricistronic vector were the *S. pombe med20/med18/med8* genes. The alignment of these and other Mediator proteins over different species revealed a high number of potential long loops in the *S. cerevisiae* proteins. This fact made it worth switching to the proteins from *S. pombe*, where secondary structure prediction points to fewer unordered regions. As a consequence of experiences with the stability of MED8, it was chosen as the tagged protein. This allows for selection of homogeneous trimeric complex. The cloned *med20* gene corresponds to the one suggested by the literature (Boube *et al.*, 2002) even though it is not clear whether a MED20 (Srb2) homolog exists in *S. pombe*, since sequence homologies are very low. When the tricistronic *med20/med18/med8*-His₆ construct is expressed in *E. coli*, MED20 does not copurify with the complex posing the question whether the used gene really codes for the correct MED20 protein and whether a MED20 homolog even exists in *S. pombe*. Further investigations will have to address the question of a different behavior and possibly different interaction pattern of MED20 in *S. pombe*.

As an indirect result of the missing MED20 it becomes clear that in *S. pombe* MED18 directly interacts with MED8. MED18 is stoichiometrically copurified with MED8 after an affinity purification step directed against MED8-His₆. This dimeric complex crystallizes (L. Lariviere, unpublished data).

1.7. Coexpression of the dimeric MED6/MED17-F2 (Srb4-F2) and the trimeric MED20/MED18/MED8 (Srb2/Srb5/MED8) complex results in a pentameric head module subcomplex

The head module scaffold protein MED 17 (Srb4) binds MED6 as well as MED8. MED8 in turn was shown to bind the MED20/MED18 (Srb2/Srb5) heterodimer. To reconstitute a subcomplex consisting of MED6, MED17, MED8, Med20 and MED18 from recombinant proteins, the bicistronic GST-MED17-Med6-His₆ construct was coexpressed with the tricistronic MED8-MED20-MED18-His₆ construct. A copurification via the only GST-tag in this five subunits and subsequent copurification over anion exchange and gelfiltration proves the predicted binding pattern (Fig. 10). The presence of all subunits in the complex was verified by mass spectrometry. Nevertheless, a high percentage of the purified protein elutes from a superose column in the void volume, and is thus aggregated. The individual subunits were not stoichiometrically present, which could be the result of different stabilities or impaired binding because of the GST-tag. Interestingly MED17-F2 and MED18 appear rather stoichiometric.

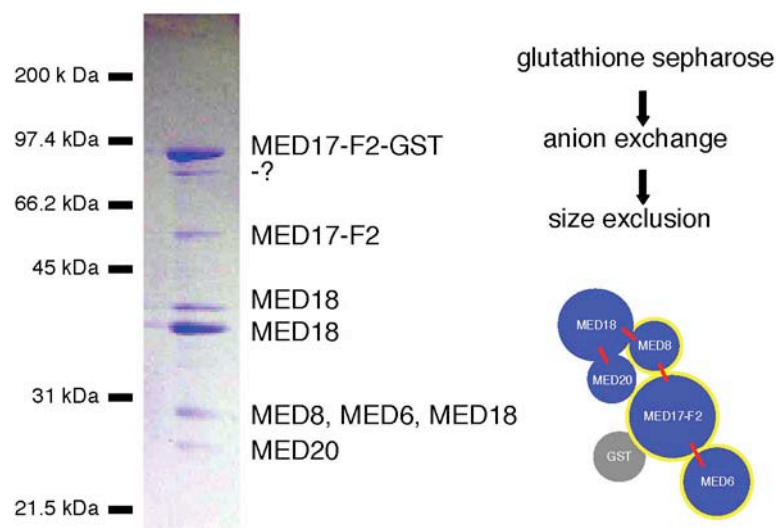


Fig. 10 – A hetero-pentameric subcomplex of the Mediator head module is purified after coexpression of bi- and tricistronic constructs. The GST-MED17-F2/MED6 Δ C construct was coexpressed with the MED8/MED20/MED18-His construct and purified over a glutathione sepharose, an anion exchange and a size exclusion column. Resulting bands were cut from a Coomassie stained gel and analyzed by mass spectrometry. All the originally expressed subunits or their degradation product were purified using just a single affinity tag on MED17 (SRB4), consistent with the predicted interaction map.

1.8. Overview on discovered contacts in the Mediator head module

Figure 11 represents an interaction map of the Mediator head module as it was derived from this work. These data are consistent with interaction studies using yeast two hybrid screens and pulldown assays (Guglielmi *et al.*, 2004; Ito *et al.*, 2001; Kang *et al.*, 2001; Koh *et al.*, 1998; Lee *et al.*, 1998; Uetz *et al.*, 2000). The stoichiometric copurification to a high degree of purity in big amounts after affinity tagging of a single subunit is the convincing proof for such an interaction. MED6 is so far the only known bridge to the Mediator middle module (Baumli *et al.*, 2005).

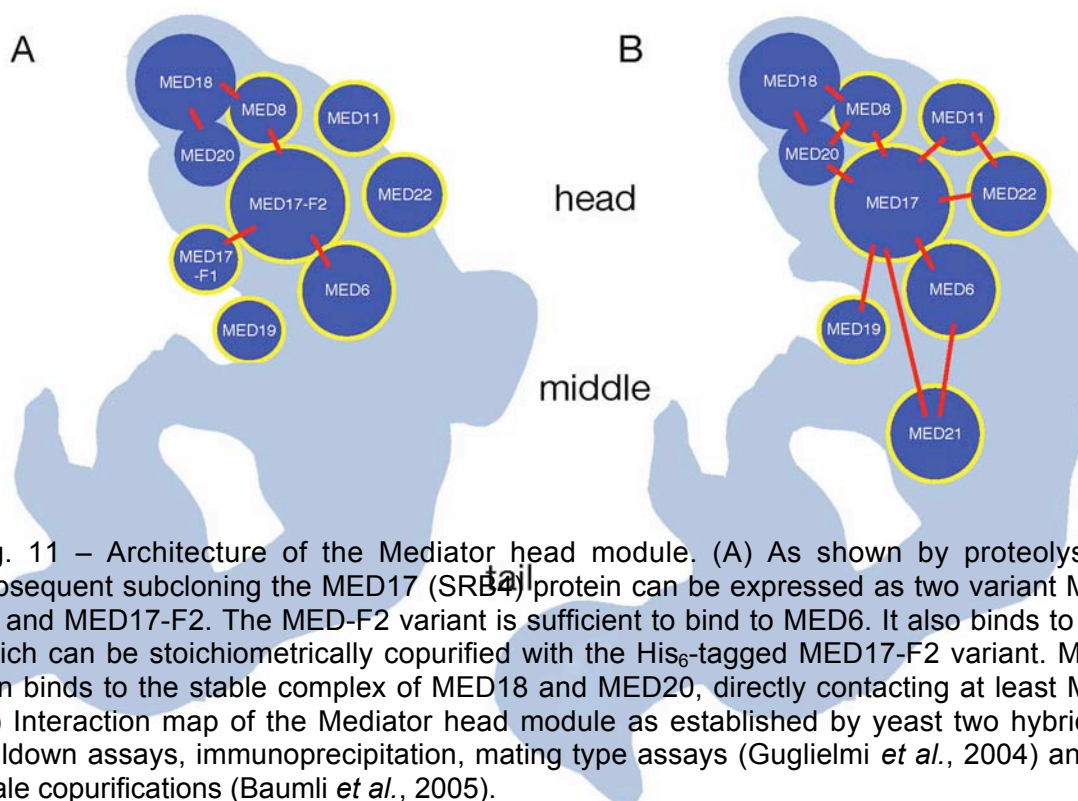


Fig. 11 – Architecture of the Mediator head module. (A) As shown by proteolysis and subsequent subcloning the MED17 (SRB4) protein can be expressed as two variant MED17-F1 and MED17-F2. The MED-F2 variant is sufficient to bind to MED6. It also binds to MED8, which can be stoichiometrically copurified with the His₆-tagged MED17-F2 variant. MED8 in turn binds to the stable complex of MED18 and MED20, directly contacting at least MED18. (B) Interaction map of the Mediator head module as established by yeast two hybrid GST-pulldown assays, immunoprecipitation, mating type assays (Guglielmi *et al.*, 2004) and large scale copurifications (Baumli *et al.*, 2005).

1.9. Binding assays point towards interactions between Mediator subcomplexes and RNA Pol II

According to EM analysis the head is the major RNA Pol II interacting module of Mediator. To test whether recombinant subcomplexes obtained in this work bind to RNA Pol II, an excess of purified GST-MED17-F2/MED6 Δ C-His (Fig. 12A) as well MED18-His₆/MED8/MED20 (Fig. 12B) was incubated with RNA Pol II (lacking Rpb4 and Rpb7) and subsequently subjected to gelfiltration. While free subcomplexes elute in a separate peak from the column, some protein is still found co-eluting with RNA Pol II, indicating a weak interaction. In the case of GST-MED17-F2/MED6 Δ C-His it

cannot be excluded that RNA Pol II binds unspecifically to GST. In addition the heterodimeric complex has a great tendency to form aggregates, which elute in the void volume. Thus the presence of MED17 (Srb4) in nearly every fraction would need further control experiments. Results from individual MED17-F2 purifications indicated that high salt concentrations in the purification buffer help to prevent aggregation of MED17-F2.

A minor fraction of the heterotrimeric MED18/MED8/MED20 complex coelutes with RNA Pol II and does not form any aggregates. The substoichiometric presence of the trimer in the RNA Pol II elution peak points to a low affinity to RNA Pol II. Nevertheless it is unlikely that the presence in this elution stems from a brought elution profile of the trimeric complex since the signal of MED18 does not faint away throughout the first four fractions. To test whether the trimeric complex is bound to RNA Pol II a second gelfiltration of the RNA Pol II containing peak was performed. Silver stained gels of the elution indicated that some trimeric complex is still present (not shown), strongly indicating binding.

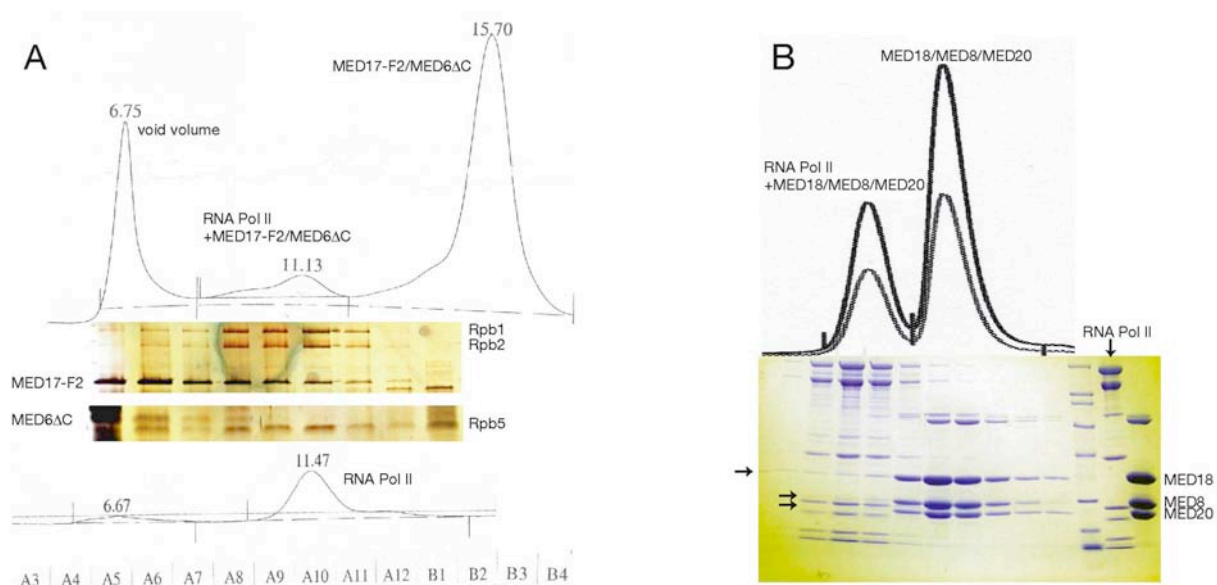


Fig. 12 – Size exclusion chromatography of RNA Pol II core-Mediator subcomplex preparations. Mediator head module subcomplexes coelute with RNA Pol II on gelfiltration columns: Superose6 elution profiles and SDS-PAGE analysis of gelfiltration fractions after binding of RNA Pol II to GST-MED17-F2/MED6 Δ C-His (A) and MED18-His₆/MED8/MED20 (B). Arrows mark the MED18, MED8 and MED20 subunits in the coeluting fractions.

2. Structural studies of the Mediator CDK8/Cyclin C module

The fourth module of Mediator, the CDK8/Cyclin C module, consists of four subunits: MED12, MED13, as well as CDK8 and its cyclin partner Cyclin C. To investigate the function of the CDK8/Cyclin C module, which is not permanently associated with Mediator, structural studies were performed on CDK8 and Cyclin C.

2.1. Cyclin C and CDK8 from *S. pombe* are obtained by recombinant expression techniques

2.1.1. BLAST search discovery and cloning of Cyclin C and CDK8

Sequences from the *S. cerevisiae* proteins revealed long loops in an alignment when compared with other species. Advantage was again taken from the fact that *S. pombe* proteins often comprise fewer regions of low complexity, which might result in a higher success rate for recombinant expression. The *S. pombe* homologs of the CDK8/Cyclin C module subunits were not known at the beginning of this project. A BLAST search determined potential candidates for Cyclin C (Fig. 13) and CDK8. The orf numbers SPBC12D12.06 (Cyclin C) and SPAC23H4.17c (CDK8) were cloned from *S. pombe* cDNA library.

```
>UNIPROT:SRB11_SCHPO O94503 RNA polymerase II holoenzyme cyclin-like subunit
(Suppressor of RNA polymerase B srb11).
Length = 228

Score = 307 (113.1 bits), Expect = 1.1e-26, P = 1.1e-26
Identities = 63/163 (38%), Positives = 99/163 (60%)

Query:   76 IYCYFLIMKLGRRNLNIRQYALATAHIYLSRFLIKASVRE-INLYMLVTTTCVYLACKVEEC 134
          IY + ++   G RL +RQ  LATA + L R+++K + + +L  LV TC+YL+CKVEEC
Sbjct:   33 IYQWKVVQTFGDRRLRLRQVRLATAIVLLRRYMLKKNEEKGFSLALVATCIYLSCKVEEC 92

Query:   135 PQYIRTLVSEARTLWPEFIPDPPTKVTEFEFYLLLEELESYLIVHHPYQSLKQIVQVLKQP 194
          P +IRT+ +EA LW  +  + ++E EF ++  L+++LIVHHPY SL+Q
Sbjct:   93 PVHIRTICNEANDLWLSKVKLSRSNISEIEFEIISVLDAPLIVHHPYTSLEQAFH----- 147

Query:   195 PFQITLSSDDLQNCWLSLINDSYINDVHLLYPHIIAACLFI 237
          ++  L+ WS++NDSY + + L+  PH +A A L I+
Sbjct:   148 --DGIINQKQLEFAWSIVNDSYASSLCLMAHPHQLAYAAALLIS 188
```

Fig 13 – result of a BLAST search identifying a potential homolog of Cyclin C in *S. pombe*. The WU-BLAST (www.ebi.ac.uk) was run with default parameters.

At a later stage, they were published as the Cyclin C and CDK8 homologs of *S. pombe* (Samuelson *et al.*, 2003). Initially Cyclin C was cloned in a bicistronic vector with CDK8 as well as in a monocistronic vector, the tag each time being

located on the C-terminal end of Cyclin C. In the bicistronic system CDK8 was expressed but was insoluble and would not bind to Cyclin C, which was well expressed. Therefore a different strategy for CDK8 expression was set up.

2.1.2. CDK8 expression with the baculovirus *Bac-to-Bac* system

Apart from missing binding partners, another reason for eukaryotic proteins to be insoluble or non-functional when expressed in bacterial cells can be missing post-translational modifications. In most cases, bacterial cells are not able to provide the eukaryotic protein with correct phosphorylations, glycosylations etc. In these cases overexpression in a eukaryotic system is advisable. The principle of the baculoviral system is the infection of insect cells with an insect cell specific virus where the gene of interest was cloned behind the strong promoter of a gene nonessential for this virus in cell culture.

For subsequent studies on CDK/cyclin interaction and crystallization trials the expression of CDK8 from *S. pombe* with the help of the baculoviral system *Bac-to-Bac* (Invitrogen) was established. The gene was cloned in the pFastBac1 vector and an additional C-terminal His₆-tag was introduced via the 5' primer. In a subsequent transposon mutagenesis step combined with a classical blue-white selection, the clone carrying the recombination event in the Dh10Bac strain was isolated. This bacterial strain carries the viral genome on a so-called Bacmid of >135 kb. In the positive clone, a big white colony on a selective blue-white medium, the CDK8 gene was fused to the strong polyhedrin promoter and flanked by the transposon sites. This step was controlled by PCR techniques using M13 primers and by CDK8 specific primers.

Production of the Bacmid DNA is a highly critical step in production of the Baculovirus. The DNA needs to be very clean and highly concentrated. Handling of this DNA has to be performed very carefully not to break it by shearing forces. Manual DNA preparation techniques, as suggested by the *Bac-to-Bac* manual (Invitrogen), did not lead to virus production, probably due to low DNA concentrations. An adapted protocol for a Qiagen Midi Prep Kit helped to solve the problem. In order to increase efficiency for elution of high molecular weight DNA, the elution buffer needs to be heated to 65°C.

For virus production in insect cells the recombinant Bacmid DNA was transfected into SF9 insect cells (Novagen). The DNA to Lipid ration optimal for this transfection was determined by evaluating infection efficiency. After the DNA is transfected into the insect cells, the recombinant virus starts its production and the expression of the protein of interest. The virus was harvested from the supernatant and amplified in sequential rounds of virus infection, in which and multiplicity of infection (MOI) were optimized stepwise. In figure 14 the establishment of the CDK8 expression technique is summarized and crucial steps are outlined.

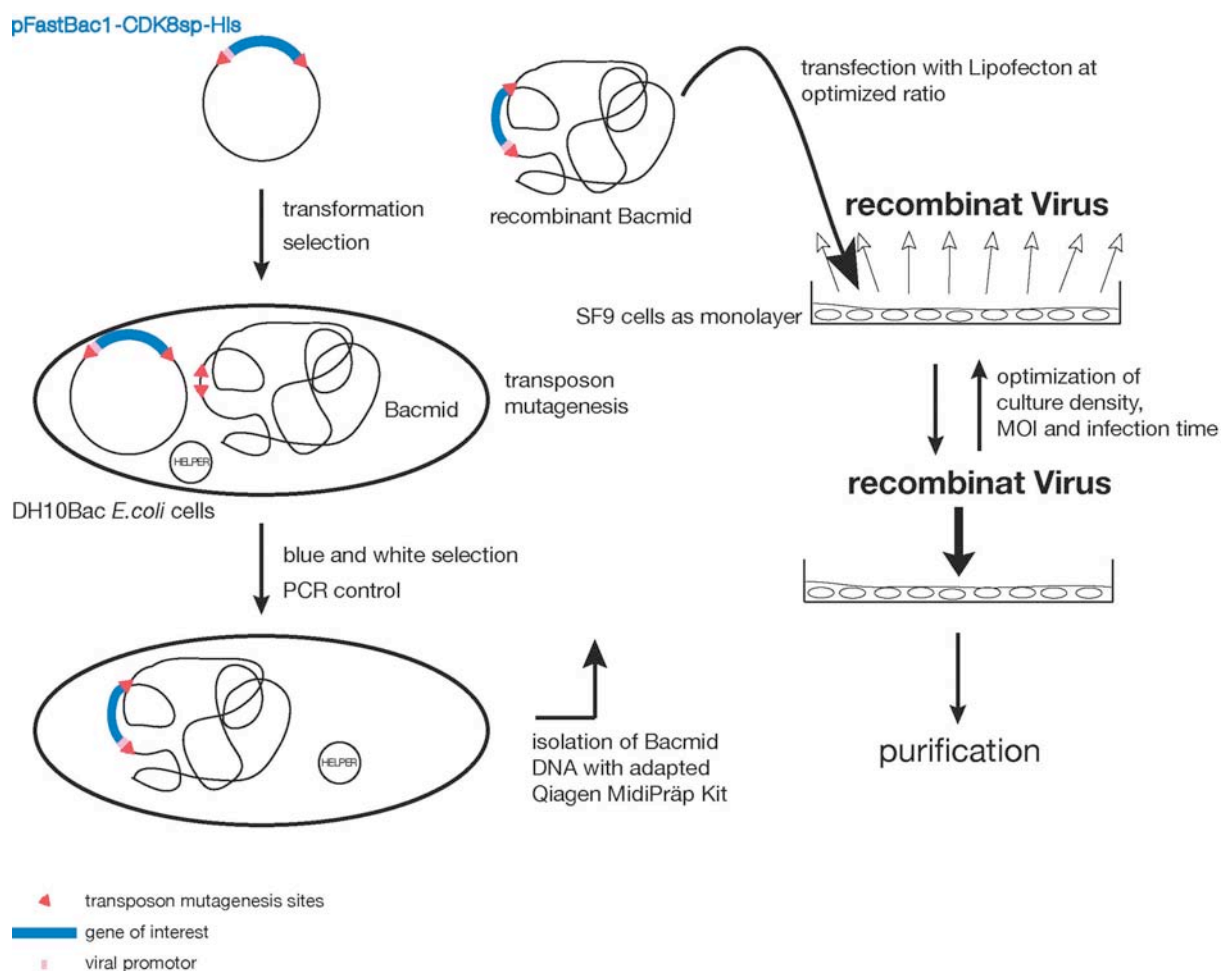


Fig. 14 – Summary of the cloning and expression procedure for recombinant production of CDK8 in SF9 insect cells. The CDK8 gene of *S. pombe* was cloned in the pFastBac1 transfer vector before transforming this vector to DH10Bac *E. coli* cells carrying the viral genome on a so-called Bacmid. After performing a blue/white selection the recombination event was verified by PCR and the Bacmid DNA was isolated via an adapted protocol with a Qiagen kit. For transfection into SF9 cells Lipofectin (Invitrogen) was tested for transfection efficiency and an optimal DNA:Lipid ratio was established. Protein production after infection with the recombinant virus was optimized by culture density, MOI, and infection time.

Expression of CDK8 was tested with 20 ml cell culture experiments. Small-scale Ni-NTA bead pulldown assays were analyzed by Western Blot for the best MOI and infection time. For protein production 500 ml of log phase 80% confluent insect cells were infected at a 1:7 ratio from the produced virus stock. Cells were harvested 48h post infection when production of the protein versus dead and broken cells was optimal. A purification protocol for CDK8 was established consisting in sonication lysis, Ni-NTA affinity chromatography and cation exchange column (Fig. 15). After optimization steps the purification of CDK8 was promising but would still need further improvement in order to reach yields suitable for crystallization.

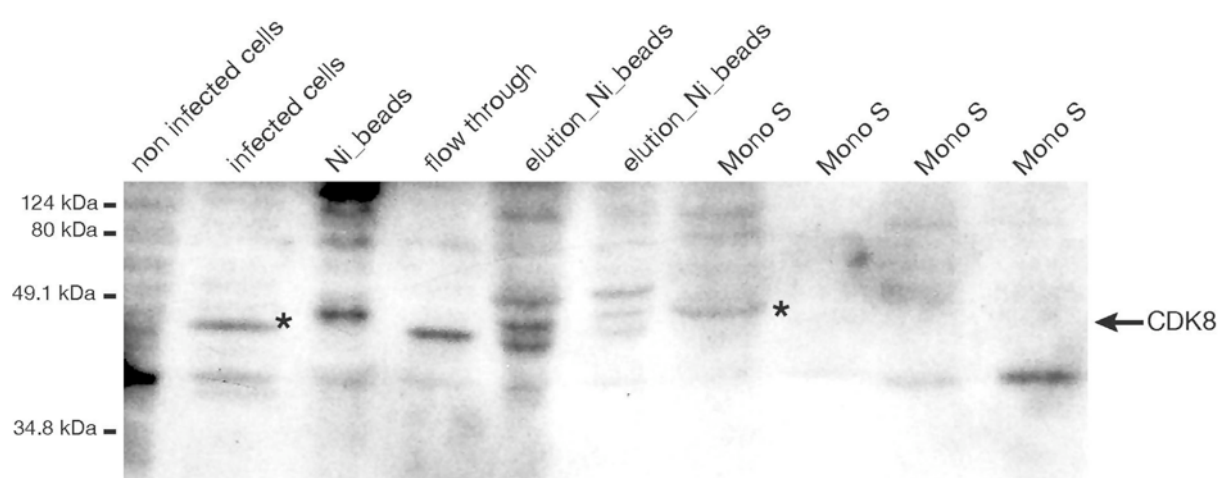


Fig. 15 – Purification of CDK8. Infected insect cells were harvested 48 h post infection, lysed by sonication. Cleared lysate was purified over a Ni-NTA column, elution fractions were pooled and subjected to a Mono S column. Uninfected control and purification fractions were analyzed by Western Blot with an anti-His antibody. The CDK8 band is marked with star in the lysate from infected cells and the Mono S fraction.

2.1.3. Cyclin C can be recombinantly expressed in *E. coli* and purified to crystallizable amounts

When cloned with a C-terminal His₆-tag, expression of full-length *S. pombe* Cyclin C yielded reasonable amounts of protein (for purification procedure compare Fig. 16 and Materials and Methods). However it was very prone to degradation and aggregation and could only be used for crystallization during two days. Cyclin C formed small fragile crystals in 200mM MgFormate that were refined to approx. 70 x 70 x 70 μm^3 diameter when grown in a buffer containing 25% glycerol, the minimal concentration needed in the cryo-solution. The Cyclin C crystals only diffracted to a maximum of 3.5 Å resolution with synchrotron radiation when they were harvested in

their cryo-solution and plunged into liquid nitrogen within seconds. Molecular replacement with these data failed. Since the crystals were so sensitive to any kind of stress, it was suspected that some secondary structure in the protein would still interfere with crystallization. To improve crystal quality, the protein was subjected to limited proteolysis.

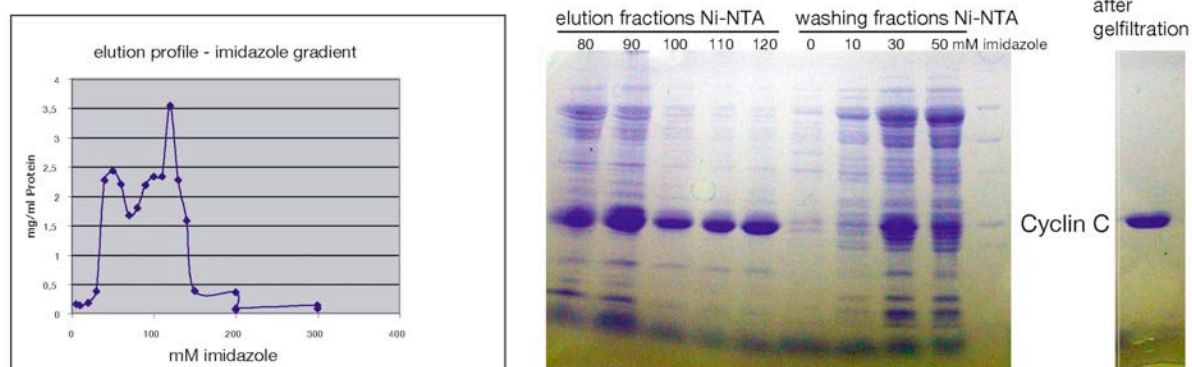


Fig. 16 - Purification of Cyclin C. Elution from the Ni-NTA column was optimized by an imidazole gradient in the elution buffer. The peak containing the most pure Cyclin C (lane 1-5) was applied to a Mono Q column before gelfiltration.

2.1.3.1. Limited proteolysis - the protease can make the difference

When testing proteins for stable domains via limited proteolysis several proteases with different specificities should be used and compared. In case of the Cyclin C using chymotrypsin or trypsin resulted in no defined proteolysis products (Fig. 17, left panel). The protein cumulatively degrades over time without stable intermediates suitable for subcloning. In contrast to that, subtilisin and proteinase K treatment resulted in cleavage of four or 29 residues from the cyclin N-terminus (Fig. 17, right panel). A protein variant that lacked the N-terminal 29 residues was insoluble, but a variant truncated by four amino acids produced crystals of improved stability that diffracted to 3.0 Å resolution (Fig. 17). Molecular replacement with known cyclin structures again failed.

Therefore the structure had to be solved by MAD analysis with a selenomethionine-substituted crystal of a mutant protein that carried replacements of two leucines to methionines. Using native diffraction data to 3.0 Å resolution, the structure was refined to a free R-factor of 28.8% (Table 24). The final model shows excellent stereochemistry and comprises all residues of Cyclin C except residues 1-5 (Table 24).

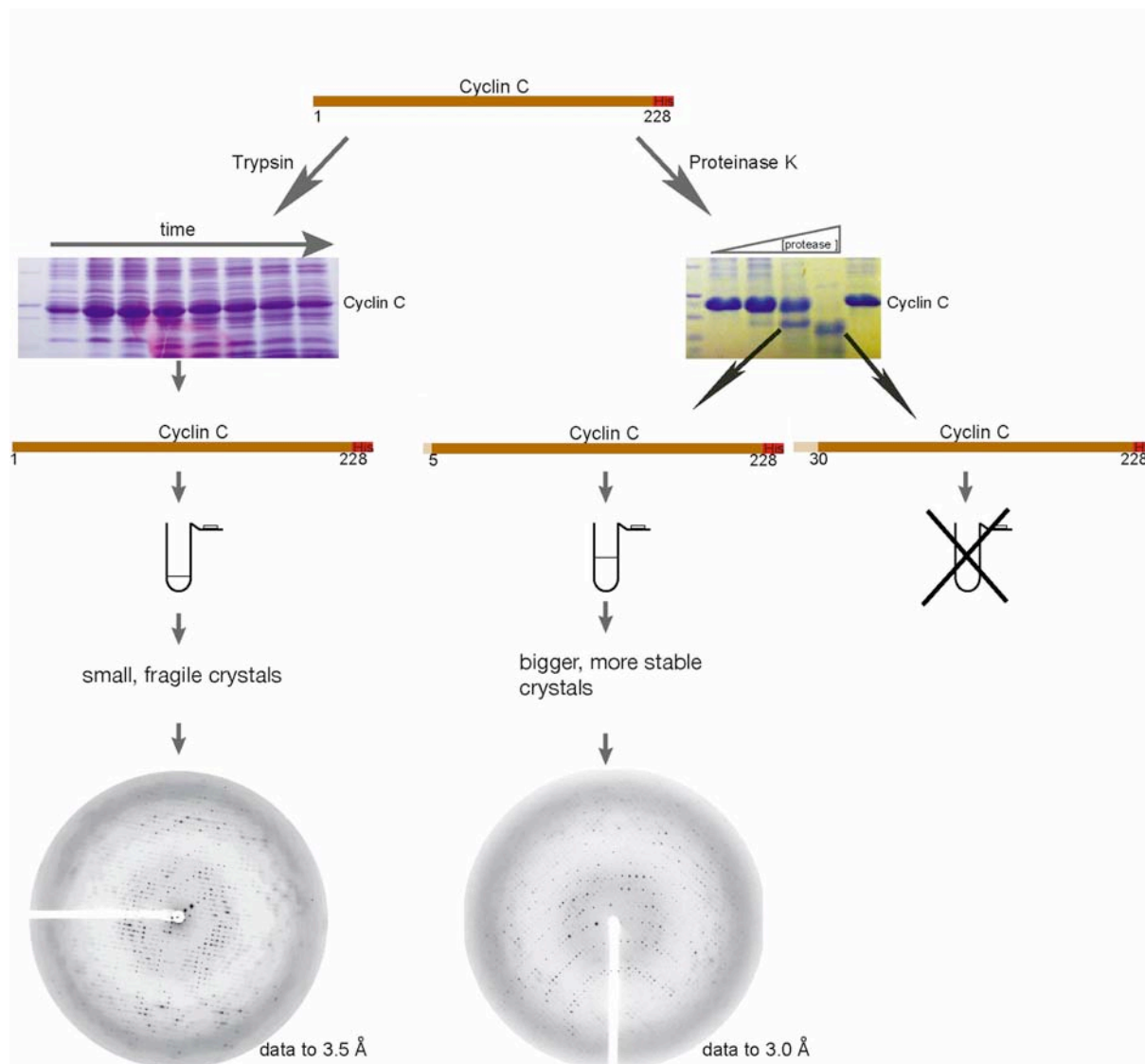


Fig. 17 – Four amino acids on the N-terminus of Cyclin C interfere with crystal quality. Full-length Cyclin C was subjected to proteolysis with trypsin and with proteinaseK. The tryptic digestion did not reveal stable fragment, the protein appeared stable over time. Crystallization of full-length Cyclin C resulted in fragile, weakly diffracting crystals. In contrast, proteinase K digestion pointed to two potential variant, only the one lacking just the first 4 amino acids resulted in soluble protein. This variant resulted bigger, more stable crystals that diffracted to 3.0 Å.

2.1.3.2. Two leucines in Cyclin C were mutated to methionines to allow MAD phasing

When molecular replacement with known structures is impossible, the solution of the phase problem can either be achieved from crystals derivatized with heavy metals or crystals from proteins that contain a heavy atom incorporated through an unusual amino acid. Selenomethionine can be used as such a variant heavy amino acid (Budisa *et al.*, 1995). For incorporation, methionine auxotroph bacteria are grown in a minimal medium with selenomethionine forcing them to incorporate the

amino acid in any protein they express. The 25 kD Cyclin C contains three natural methionines including the N-terminal starting methionine. Ideally, methionines should be located in the hydrophobic core because this augments the chances that they are well ordered in the crystal. Since two methionines might not result in sufficient phasing power, three additional targets for mutations were chosen: L60M, I125M, and L186M.

Subsequently, the protein sequence was mutated using two-step PCR setups. The resulting triple mutant unexpectedly yielded only insoluble protein. Testing all combinations of two mutations a variant with an L60M and an L186M substitution was found to be soluble. Interestingly, insolubility was caused only by the isoleucine to methionine mutation, even though this isoleucine is highly conserved and located in the hydrophobic core of previously solved cyclin structures.

The purification procedure for the mutant protein was adjusted since a different elution profile from affinity chromatography columns was observed. Expression levels in the methionine auxotroph strain in minimal medium were much lower and the protein was very unstable. It resulted within a day to crystals of up to $150 \times 150 \times 150 \mu\text{m}^3$. Diffraction data sets with a maximum resolution of 3.5 Å were measured with synchrotron radiation.

2.2. Solution of the Mediator Cyclin C structure

2.2.1. Data collection

Due to their limited size, low diffracting power and high solvent content crystals had to be measured and screened at synchrotron radiation. Using the home source (rotating anode, Rigaku), hardly any diffraction was visible. Probably due to their high solvent content (72%), crystals were also very sensitive to harsh changes and osmotic stress. Additionally they were very radiation sensitive. To optimize the data collection strategy, MOSFLM (Powell, 1999) was routinely used to set up the strategy. A highly redundant and complete data set needed around 100° . To avoid local high x-ray doses that would trigger premature decay of the crystal, it was absolutely crucial, to focus the beam on the detector instead of the crystal. Data sets that were collected with a different setting were not complete, due to early crystal damage. For phasing a MAD experiment was carried out around the selenium edge,

collecting a 100° datasets first at the peak and subsequently at inflection and remote wavelength (0.97977 Å, 0.98004 Å, and 0.94927 Å respectively).

2.2.2. Phasing and refinement

Molecular replacement with the program AmoRe failed with data sets from full-length and mutant Cyclin C crystals. Cyclin structures might deviate substantially outside the cyclin repeats and thus a search model was designed to use in molecular replacement. To this end known cyclin structures from Cyclin A, Cyclin H and Cyclin V were superimposed and only the core regions that possess a very similar backbone fold were transferred into a new pdb-file. Coordinates were taken from Cyclin H which shows the highest degree of sequence identity to Cyclin C. All non-identical (Cyclin C to Cyclin H) residues were mutated to alanines. An alternative search model consisted only of alanines. Failure of the program to phase the data set even with the help of these search models was probably due to the very limited resolution of the data, combined with a rather different topology of the molecule in question.

Phasing was achieved with the MAD experiment. Data from peak, inflection and remote were processed and phases could be obtained with the help of the SOLVE program. The peaks for all four ordered methionines – except for the N-terminal methionine – could be located and served as a sequence marker for model building. The “consensus cyclin model” that had been constructed for molecular replacement was fit into the initial electron density and phase combination using CNS helped to improve the electron densities. Due to the limited resolution the model was built by hand, including more details and improving the phases at each step. To prevent model bias the model was then truncated to a polyalanine chain and phases from this model were combined with amplitudes from the native crystal data set. Thereby the resolution was extended from 3.5 Å to 3.0 Å. This allowed building of the atomic model by hand. Refinement of the structure with CNS (Brunger *et al.*, 1998) resulted in a free R-factor of 28.8% (Table 1). In the refined structure, 98.6% of the residues fall in allowed regions of the Ramachandran plot, and none of the residues are found in disallowed regions (Laskowski *et al.*, 1993a).

Table 24 – Cyclin C structure determination and refinement

Crystal	Cyclin C (5-228) variant L60M/L186M, SeMet MAD			Wild type, native
Data collection				
Space group	P3 ₁ 21			P3 ₁ 21
Unit cell axes (Å)	A=b=91.9; c=90.9			a=b=91.7; c=90.1
Wavelength (Å)	0.97977 (peak)	0.98004 (inflection)	0.94927 (remote)	1.0725
Resolution (Å)	20-3.5 (3.6-3.5) ¹	20-3.5 (3.6-3.5)	20-3.5 (3.6-3.5)	20-3.0 (3.16-3.0)
Completeness (%)	99.4 (100)	99.5 (100)	99.7 (100)	99.4 (99.4)
Unique reflections	5866 (581)	5873 (581)	5885 (581)	9161 (1333)
Redundancy	9.12	9.04	9.29	6.0
R _{sym} (%)	8.6 (28.8)	7.4 (35.9)	5.8 (27.7)	8.1 (41.6)
Mean I/s I	23.6 (7.3)	23.8 (6.6)	24.5 (6.1)	17.0 (3.0)
<i>f</i> '	-7.0	-10.2	-2.8	-
<i>f</i> ''	5.5	3	3.6	-
Refinement				
Residues				227 ²
RMSD bonds (Å)				0.008
RMSD angles (°)				1.3
R _{cryst} (%)				24.4
R _{free} (%)				28.8

¹The numbers in parenthesis correspond to the highest resolution shell.

²The final model comprises residues 6-228 of Cyclin C, and four additional residues at the C-terminus that have been introduced during cloning as a spacer to the hexahistidine tag.

2.3. Analysis of the Cyclin C structure

2.3.1. The structure of Cyclin C - canonical cyclin repeats

The structure of Cyclin C reveals two typical cyclin repeats consisting of five helices each (H1-H5 and H1'-H5' in repeat 1 and 2, respectively) (Fig. 18). The length of some helices differs from that in cyclins A and H by several residues (Fig. 18). The individual cyclin repeats 1 and 2 can be superimposed on those of Cyclin H with a root mean square deviation (RMSD) in main chain atom positions of 1.6 Å and 2.5 Å, respectively. The relative orientation of the two cyclin repeats is also similar to that observed in cyclins A and H, so that both repeats of Cyclin C can be superimposed with a RMSD of 2.7 Å in each case. The similar orientation of the two repeats in cyclins C, H and A apparently results from conservation of the repeat interface, which involves residues from helices H1, H2, H1', H2', and the conserved inter-repeat linker (Fig. 18). In particular, the highly conserved residue R62 in H2 binds to the linker backbone carbonyl of residue F132 in the inter-repeat linker, as observed in Cyclin H (Andersen *et al.*, 1996a; Lolli *et al.*, 2004), and additionally the carbonyl of I134. In addition, residue Q50 in helix H1, Y139, and N164 in H2' form a cluster at the repeat interface. In contrast, the two cyclin repeats in the general transcription factor TFIIIB adopt a different relative orientation (Nikolov *et al.*, 1995).

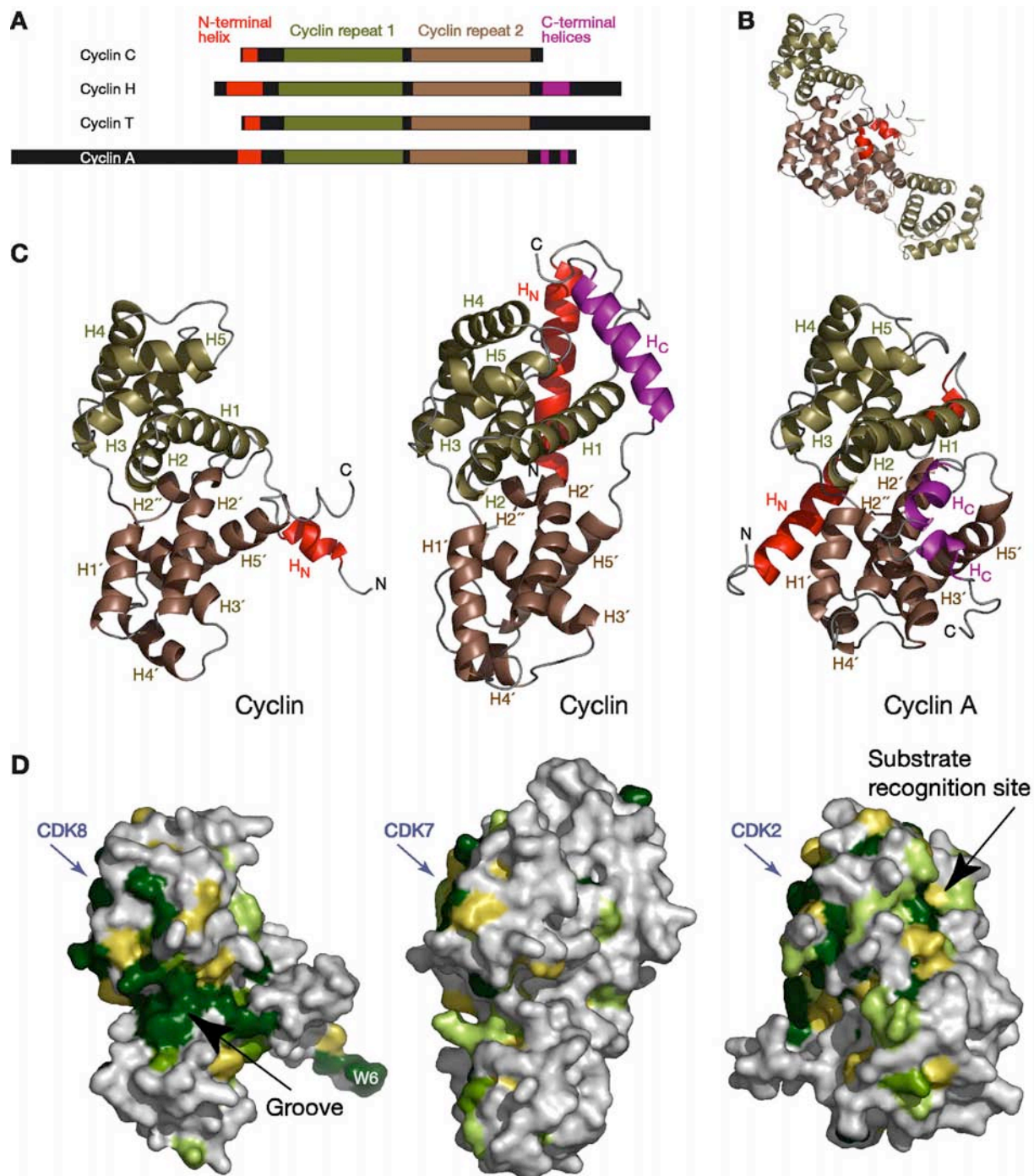


Fig. 18 – Structure of Cyclin C and comparison with cyclins A and H. A, Schematic diagram of cyclin primary structures. The two canonical cyclin repeats and the N- and C-terminal helices are highlighted. Bars are drawn to scale. B, Ribbon model of the crystallographic dimer of *S. pombe* Cyclin C. C, Structural comparison of cyclins C, H (Andersen *et al.*, 1996a) (PDB-code 1JKW), and A (Jeffrey *et al.*, 1995) (PDB-code 1FIN). The proteins are drawn as ribbon models. The color code for secondary structure elements is as in A and in figure 19. Helices are numbered according to figure 19. D, Surface conservation. The surface of the three cyclins in C is colored according to conservation as indicated in figure 19. Residues are highlighted in dark green, green, light green, and yellow, according to decreasing degree of conservation. The view is as in C. Figures prepared with PYMOL (DeLano Scientific).

2.3.2. A mobile N-terminal helix

Cyclin C differs from known cyclin structures mainly outside the canonical repeats. Before the first repeat, Cyclin C contains only one N-terminal helix (H_N), whereas Cyclin A additionally contains an extended N-terminal region (Fig. 18). The C-terminus of Cyclin C is formed by the last helix in the second repeat, whereas Cyclins H and A comprise additional C-terminal α . Cyclin C may thus be regarded as a minimal cyclin, that only consists of the two canonical repeats and a specific H_N helix.

Helix H_N in Cyclin C differs from that in Cyclins A and H in three aspects (Figs. 18C, 19). The helix is much shorter, adopts a different position, and is mobile. The H_N helix protrudes from the cyclin repeats, whereas the N-terminal helices of Cyclins H and A intimately pack on the repeats with conserved residues (Fig. 18C-D) (Andersen *et al.*, 1996a; Lolli *et al.*, 2004) The Cyclin C helix H_N and the subsequent loop show poor electron density, and have B-factors well above the average B-factor for the two cyclin repeats (133 \AA^2 on average for residues 5-29; 75 \AA^2 for residues 30-228). However the N-terminal helices in Cyclins H and A show B-factors that are comparable to the overall B-factors of the structures. Also, partial proteolysis of Cyclin H does not cleave its N-terminal helix, indicating that it is an integral part of the structure. The rigid nature of the N-terminal helices in Cyclin A, and also in a viral cyclin, is further indicated by the observation that they adopt the same position in the free cyclins and in their complexes with CDK2 (Card *et al.*, 2000; Jeffrey *et al.*, 1995).

2.3.3. Structure-based alignments.

Using the structures of Cyclin C (this work), Cyclin H (Andersen *et al.*, 1996a; Lolli *et al.*, 2004) and Cyclin A (Card *et al.*, 2000; Jeffrey *et al.*, 1995) a structure-based alignment of these different cyclin families using sequences from *S. pombe*, *S. cerevisiae*, and human was prepared (Fig. 19). The structure-based alignment allows identification of residues that only occur in a certain type of cyclin and are strictly conserved there, and thus are candidates for cyclin-specific functions (see below). The sequence comparisons also revealed that Cyclin C is the most conserved of all cyclins. Human Cyclin C shares 71% and 29% of its residues with Cyclin C from *Drosophila melanogaster* and *S. cerevisiae*, respectively, whereas the corresponding percentages of identical residues are 32% and 21% in Cyclin A, and 40% and 26% in

Cyclin H, respectively. In particular, the conservation of the second cyclin repeat is much more extended in Cyclin C compared to other cyclins.

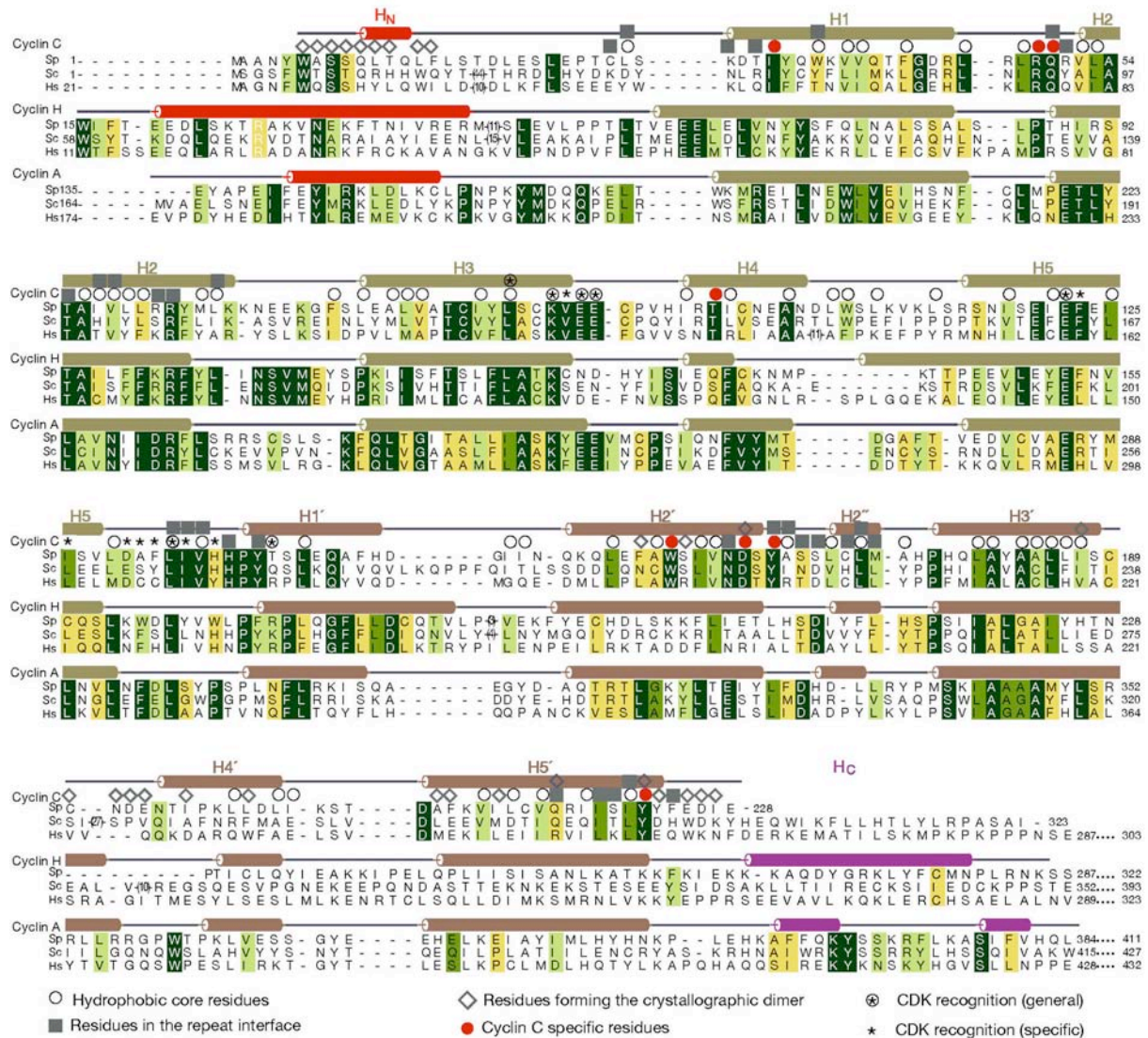


Fig. 19 – Structure-based sequence alignment of cyclin families C, H, and A. Structure-based alignment of protein sequences of *S. pombe* Cyclin C (this study) with human Cyclin H (Andersen *et al.*, 1996a; Lolli *et al.*, 2004), and human Cyclin A (Jeffrey *et al.*, 1995). Alignments were extended with Clustal W (Thompson *et al.*, 1994) to three members of each cyclin family, from *S. pombe* (Sp), *S. cerevisiae* (Sc), and *H. sapiens* (Hs), and were edited by hand. Helical regions were determined with DSSP (Kabsch and Sander, 1983) and are shown above the alignments as cylinders. Residues are highlighted in dark green, green, light green, and yellow, according to decreasing degree of conservation. Residues in the structural core, the repeat interface, and the crystallographic dimer contact are marked with open circles, filled squares, and open diamonds, respectively. Residues that are strictly conserved and at the same time specific for Cyclin C are marked with a red dot. Residues involved in CDK recognition are marked with asterisks (compare table 25). Prepared with ALSRIPT (Barton, 1993).

2.3.4. Modeling of the CDK8-Cyclin C pair.

A model of the CDK8-Cyclin C complex could be obtained in two steps. First, Cyclin A in the human CDK2-Cyclin A complex crystal structure (Jeffrey *et al.*, 1995) (PDB-code 1FIN) was replaced with the *S. pombe* Cyclin C structure. To this end, residues in Cyclin C that are near the CDK interface (residues 85-89, 94-98, and 118-129) were superposed onto their counterparts in Cyclin A (residues 263-267, 273-277, and 291-302, respectively). Second, a homology model of CDK8 was constructed by replacing all differing CDK2 residues with program O (Jones *et al.*, 1991). Homology modeling was enabled by the high degree of sequence identity between *S. pombe* CDK8 and human CDK2 (37%).

2.3.5. Aligning CDK classes within the protein family allows determination of CDK8 characteristics and their mapping on the CDK8 model

To determine specific features of the CDK8 class, a ClustalW alignment of 17 CDK sequences with two to four belonging to one class was performed. This approach is practicable due to the high degrees of sequence identity and extremely similar backbone fold that has been observed for CDKs in general (Lolli *et al.*, 2004; Russo *et al.*, 1996; Tarricone *et al.*, 2001). Similar to the “two dimensional alignment” performed for the cyclin structures, this approach allows identification of residues that are either identical in all analyzed CDKs, highly conserved among CDKs or highly conserved (identical) only in one class of CDKs while not being present at this position in other CDKs (Fig. 20). Projecting the identified residues from the alignment onto the 3 dimensional model gives insight into the nature of these specificities. Three of the residues cluster at the interface with the cyclin allowing speculations about the nature of CDK-cyclin specificity (see below). Other residues cluster on the opposite site of the kinase in an area that is known to interact with inhibitors. Clustering of specific residues in patches can always indicate potential interaction sites with binding partners important for complex formation and/or regulation. In addition analysis showed that CDK8 contains two conserved insertions, absent from other CDKs. The localization of the 9 amino acid insertion close to the activation loop and the specific aspartate has point to an important function in CDK8 activity. With

Fig. 20 (previous page) – Determination of class specific characteristics within the CDK family. With the Clustal W algorithm (www.ebi.ac.uk) CDK7, CDK8, CDK9 and CDK2/CDK3 were aligned using sequences from *S. cerevisiae*, *S. pombe*, *D. melanogaster* and *H. sapiens* at the same time. Thereby residues that are identical in all classes (blue), can be distinguished from residues conserved among one class of CDK, but also present in other classes (green) and residues that are specific and conserved only in one class of CDK (red). Similarities within one class are marked in yellow. This analysis allowed the detection of several CDK8 specific residues, among them several specific and highly conserved insertions.

2.3.6. Specificity of CDK-cyclin interaction.

Comparison of the CDK-cyclin interfaces in the CDK8-Cyclin C model and the CDK2-Cyclin A structure identified several highly conserved residues that are generally involved in CDK-cyclin interactions (“general,” Table 25, Fig. 21). Five invariant CDK residues are involved in these generally conserved interactions with cyclin residues (V54, I59, R60, R160, L170 in *S. pombe* CDK8, Table 3), consistent with previous observations (Andersen *et al.*, 1996a; Jeffrey *et al.*, 1995; Lolli *et al.*, 2004). These general interactions explain why Cyclin C can interact with both CDK8 and CDK3 (Ren and Rollins, 2004). In addition, residues in the CDK-cyclin interface that are conserved only within a certain CDK or cyclin family were identified. These “specific” residues may govern preferential interaction of a certain cyclin with a certain CDK (Table 25). The two classes of interface residues cluster in two regions, a “general” and a “specific recognition region” (turquoise and red, respectively, in Fig. 21B).

Specificity of Cyclin C for CDK8, as opposed to CDK2 or the two other CDKs involved in transcription, CDK7 and CDK9, may be explained by several contacts of side chains in the specific recognition region. In particular, several hydrophobic side chains that pack against each other in the interface of the complex model are complementary in size (Fig. 21C). For example, the side chains of L81 in the CDK8 model and F123 in the Cyclin C structure contact each other, and correspond to the contact H71-H296 in the CDK2-Cyclin A complex structure (Fig. 21C, Table 25). Residues M62 in the CDK8 model and A131 in the Cyclin C structure also contact each other, and the corresponding contact I52-F304 is observed in the CDK2-Cyclin A interface (Fig. 21C, Table 25). Similar complementary and specific pairs of residues are observed at these positions in sequences of the pairs CDK7-Cyclin H and CDK9-Cyclin T (Fig. 21, not shown).

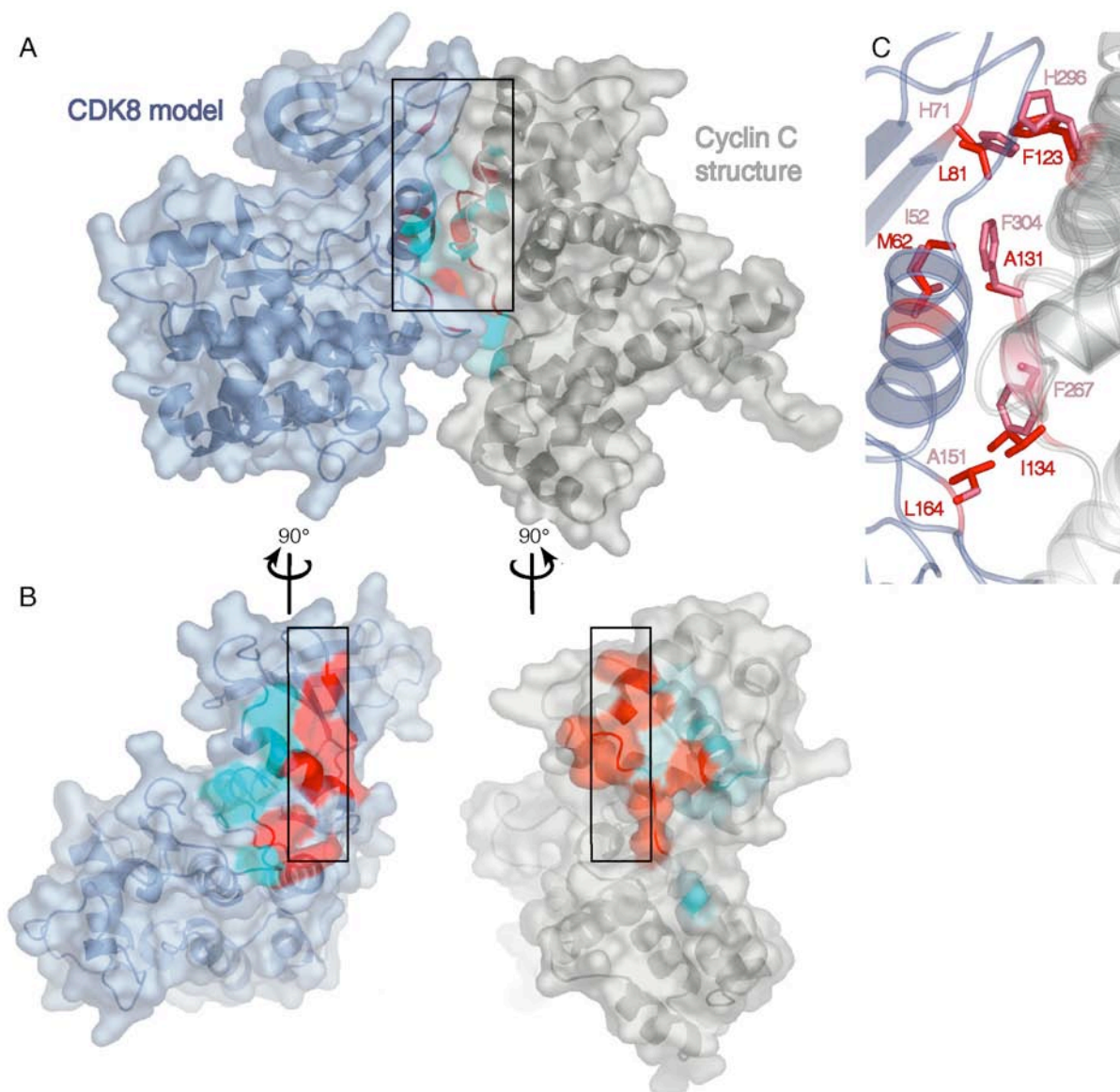


Fig. 21 – CDK-cyclin recognition. A, Model of the CDK8-Cyclin C complex, based on the CDK2-Cyclin A structure (Jeffrey *et al.*, 1995) (PDB-code 1FIN, compare text). The proteins are represented as ribbon models with a superposed semitransparent molecular surface. The black box outlines the region of the complex depicted in C. B, “Book view” of the model in A. The two protein representations were rotated by 90 degrees in opposite directions to allow viewing of the interface. Surfaces of residues that constitute the “general” and the “specific recognition region” of the CDK-cyclin interface were highlighted in turquoise and red, respectively (compare text and Table 2). C, Comparison of selected residues in the specific recognition region of two CDK-cyclin interfaces. Residues of the CDK2-Cyclin A structure (PDB-code 1FIN) are in light pink, and residues in the CDK8-Cyclin C model are in red.

This analysis however does not account for possible alternative interactions and rearrangements in the CDK-cyclin interface, which may be crucial to optimize the surface complementarity. For example, cyclins A and C apparently form alternative

salt bridges at the interface with their cognate CDKs. A conserved basic residue in the kinase (K56 in CDK2, R66 in CDK8) is bound by the non-conserved aspartate D305 in Cyclin A, but may be bound by D130 in Cyclin C, which is offset in register by two residues (Table 25). Indeed side chain rearrangements in a CDK-cyclin interface have been observed in the recent structure of a CDK2/cyclin E1 complex (Honda *et al.*, 2005), in which K56 of CDK2 adopts a different conformation and does not to interact with the basic residue R225 in cyclin E.

Table 25 – CDK-cyclin interactions¹

CDK2 structure (Jeffrey <i>et al.</i> , 1995)	CDK8 model	Cyclin A structure (Jeffrey <i>et al.</i> , 1995)	Cyclin C structure
<i>General recognition region</i>			
V44 (backbone)	V54 (backbone)	K266	K88
		E295	E122
I49	I59	K266	K88
		L263	L85
		L306	L133
R50	R60	F267 (backbone)	V89 (backbone)
		K266 (backbone)	K88 (backbone)
R150	R164	F267 (backbone)	V89 (backbone)
		E268 (backbone)	E90 (backbone)
		E269 (backbone)	E91 (side chain)
V154	L170	T316	T140
		E268	E90
<i>Specific recognition region</i>			
H71	L81	H296	F123
		-	I126
I52	M62	F304	A131
S53	M63	F304 (backbone)	A131
		F267	V89
		L306	L133
		-	F132 (backbone)
K56	R66	T303	D130 ²
		F304	A131
		D305 ²	-
A151	L164	A307	I134
		F267	-
G153	R166	Q313	T140
		-	H136

¹Depicted are CDK residues that have at least one atom within 4 Å of an atom of the cognate cyclin. For the general recognition region, cyclin residues are only included if the contacts are observed in the CDK2-Cyclin A structure and in the CDK8-Cyclin C model.

²Possible alternative salt bridges see text.

2.3.7. Conserved surfaces.

Specificity for CDK-cyclin interaction may also arise from additional protein subunits that could bridge between the kinase and the cyclin, such as MED12 (Srb8) and MED13 (Srb9) in the case of the CDK8-Cyclin C pair. To detect possible interaction sites for other proteins on cyclins C, H, and A, their molecular surfaces were colored according to conservation over species (Fig. 18D). A conserved and Cyclin C-specific surface patch is found just before the mobile helix H_N, comprising the exposed aromatic residues Y5 and W6 (Fig. 18D). Whereas the N-terminal helix of Cyclin A contributes to the interface with CDK2, helix H_N in Cyclin C points away from the kinase-binding surface (Figs. 18, 21), and cannot contribute to CDK8 binding. In the crystals, the N-terminal region of one molecule contacts the second cyclin repeat of its neighbor, resulting in a crystallographic dimer that buries several hydrophobic surface residues in its interface (L186, L200, A208, F209, I212, V215, I219) (Figs. 18B, 19). The dimer may pre-exist in solution since deletion of helix H_N results in an insoluble protein, possibly because hydrophobic residues in the dimer interface are exposed (not shown). Since most of the hydrophobic residues and the H_N helix are not conserved, the dimer is apparently not physiological. In the intact CDK8/Cyclin C module, the N-terminal region of Cyclin C could therefore bind MED12 (Srb8) or MED13 (Srb9), and, given its mobility, may change its position upon interaction with the target protein.

2.3.8. A conserved groove unique to Cyclin C.

Cyclin C also shows a highly conserved surface groove between the two cyclin repeats (Fig. 18D). The corresponding region in cyclins A and H is not conserved, and the conserved groove in Cyclin C is thus likely to have a function specific for this type of cyclin. Indeed, five of the surface residues in this groove are invariant among Cyclin C family members, but are not conserved and do not occur in any other cyclin family (residues I33, R49, W160, D165, and Y167) (Fig. 19, red dots; Fig. 22A). Since the groove is near the active site and the activation segment of CDK8 in the CDK8-Cyclin C model (Fig. 22), it may bind substrates and could contribute to CDK8 specificity. Indeed, Cyclin A has a conserved surface patch that binds kinase substrates, albeit at a different location (Fig. 18D) (Schulman *et al.*, 1998).

2.3.9. Possible mechanisms of CDK8 activation.

CDKs are generally activated in two steps that include binding of the cyclin, and phosphorylation of a conserved threonine in the CDK activation segment (T160 in human CDK2) (Pavletich, 1999). Interaction of the phosphothreonine side chain with three conserved arginines (Fig. 22B) triggers a conformational change that results in full kinase activation. All CDKs carry a threonine or a serine at the phosphorylated position, except CDK8, which has a conserved aspartate (D176 in *S. pombe*). The aspartate could however mimic a phosphothreonine, because of its negative charge and similar extension, and could interact with the three arginines, which are also conserved in CDK8 (Fig. 22C). This model however requires that the activation segment of CDK8 adopts the same conformation as in CDK2, which may not be the case, because it is three residues longer, and since CDK8 contains a specific ten-residue insertion near the activation segment (Fig. 22A, purple spheres). Indeed, the activation segment of CDK7, which is also longer, adopts a different conformation (Lolli *et al.*, 2004). Its phosphorylated threonine is not coordinated by the three conserved arginines.

An alternative mechanism of CDK8 activation is suggested by the Cyclin C structure. In the CDK8-Cyclin C model, the negatively charged side chain of the exposed glutamate E91 in Cyclin C is in a position to interact with the arginine cluster in CDK8, and could in principle activate the kinase in trans (Fig. 22C). The side chain of the corresponding residue in Cyclin A (E269) points in the opposite direction, where it forms a hydrogen bond with T231 (Fig. 22B). This hydrogen bond cannot be formed in Cyclin C, because the residue corresponding to T231 is a valine in Cyclin C (V52) (Fig. 22C). Cyclin C could in principle act similar to a viral “supercyclin” that can activate CDK6 without phosphorylation of the activation segment (Schulze-Gahmen and Kim, 2002). Finally, the exposed E91 of Cyclin C could also be involved in determination of CDK8 substrate specificity, similar to a conserved glutamate in an activator protein of CDK5 (Tarricone *et al.*, 2001).

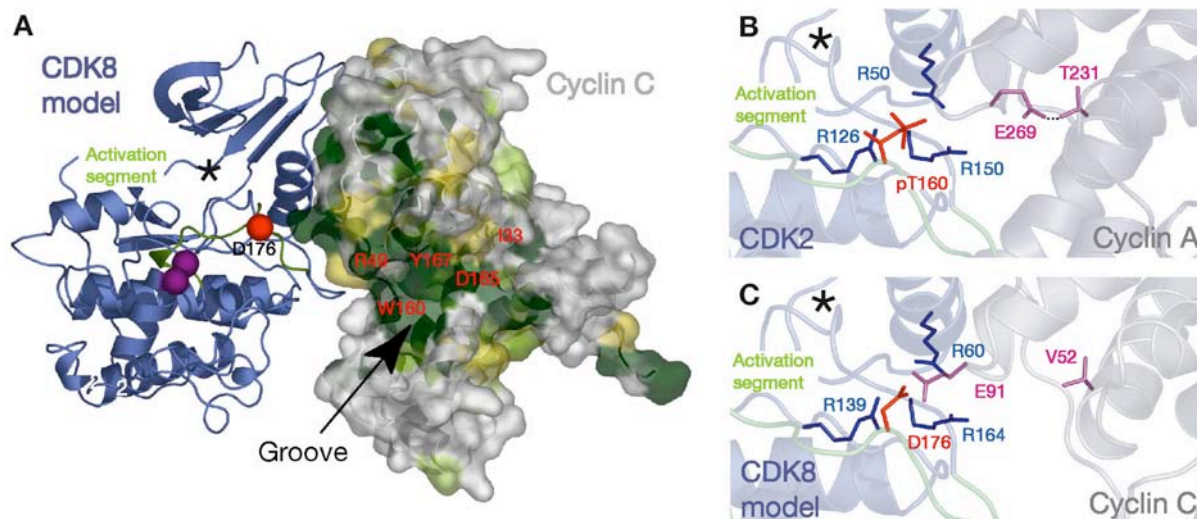


Fig. 22 – Structural features relating to CDK8 activity. A, Model of the CDK8-Cyclin C complex in figure 21A.. The model is shown as a ribbon diagram, with a semitransparent surface colored according to conservation of Cyclin C, as in figures 18, 19. The asterisk indicates the location of the kinase active site. The CDK activation segment is in green and includes the conserved aspartate D176 (red sphere). The beginning and the end point of a ten-residue CDK8-specific insertion are marked with purple spheres. The conserved surface groove in Cyclin C is indicated. B, Detailed view of the region around the activation segment in the structure of Cyclin A bound to CDK2 phosphorylated at residue T160 (Russo *et al.*, 1996) (PDB-code 1JST). The phosphothreonine is in red, and the three conserved arginines (arginine cluster) that it coordinates are in blue. A hydrogen bond between cyclin residues E269 and T231 is shown as a dashed line. C, Possible mimicry of phosphorylation in the CDK8 activation segment. The CDK8-Cyclin C model in A is oriented as in B. A conserved aspartate in the CDK8 activation segment that could mimic a phosphothreonine is in red (D176). A glutamate in Cyclin C that could bind the CDK8 arginine cluster in trans is in pink (E91). The hydrogen bond depicted in B cannot be formed, since the donor T231 in Cyclin A is replaced by a valine in Cyclin C (V52, pink).

Discussion and future directions

1. The Mediator head module: dissecting a protein-protein network

A core Mediator consists of subunits from the head and middle modules and is responsible for contacts with the general transcription machinery, RNA Pol II and GTFs. Core Mediator enables 4-fold activation of transcription in nuclear extracts, compared to 18-fold activation for the complete Mediator (Liu *et al.*, 2001; Spahr *et al.*, 2000; Spahr *et al.*, 2001). The composition of the Mediator in several species, its conservation, and the crucial role it has in transcription activation are generally accepted. Nevertheless, the molecular mechanism of this function remains enigmatic, mainly due to lack of detailed structural information. Interaction of head and middle module with RNA Pol II was shown by EM techniques, suggesting that the head module is the major RNA Pol II-interacting module.

Over the last years pull-down assays, genetic and yeast two hybrid screens, coexpression of subunits in insect cells, coimmunoprecipitation, and the split ubiquitin assay (Guglielmi *et al.*, 2004; Ito *et al.*, 2001; Kang *et al.*, 2001; Koh *et al.*, 1998; Lee *et al.*, 1998; Uetz *et al.*, 2000, Gromoller and Lehming, 2000) led to interaction maps of Mediator and the Mediator head module. With the help of bi- and tricistronic expressions and cotransformations, large-scale copurifications could be tested and thereby multiple direct protein interactions within the head Module were determined and verified. Such copurification can successfully map strong and specific direct protein-protein interactions, as demonstrated by the structure determination of the MED7/MED21 complex (Baumli *et al.*, 2005). The copurification assay with only one tagged protein 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.

1.1. MED17 – domains and interactions

Subunits of the Mediator head module are mostly insoluble when recombinantly expressed as single subunits in the *E. coli* system. The largest subunit of the head module, the essential MED17 (Srb4), had been suggested previously to

form a kind of scaffold protein in the organization of the module (Koh *et al.*, 1998). Therefore it was chosen as a starting point for structural analysis. Expression of MED17 as a single subunit yielded enough protein to start a domain analysis using the limited proteolysis approach. As a result, two variants of MED17 were designed, the F1 and F2 fragments. Secondary structure prediction, alignment studies and data available from genetic and biochemical studies (Koh *et al.*, 1998, Lee *et al.*, 1998, Thompson *et al.*, 1993) point to a great importance of the F2 fragment, since it comprises potential interaction sites with activators like Gal4, with other Mediator subunits like MED6, the *srb*-mutation as well as all sequences essential for viability (Fig. 2B). It was therefore used in a systematic bicistronic expression and purification screen with the remaining subunits of the Mediator head module. The interaction of MED17 (Srb4) with MED6 that had been observed before in a GST pulldown assay (Lee and Kim, 1998) was verified by copurification of the two subunits over two subsequent affinity chromatography steps. A very recent study had also determined this interaction in a yeast two-hybrid analysis (Guglielmi *et al.*, 2004). Successful purification of the complex was independent of the order of the affinity columns. As it could be expected from the genetic data, the weakly conserved N-terminal F1 fragment of MED17 was not required for interaction. For MED6, a C-terminal truncation mutant without the non-conserved C-terminus was used (residues 1-214) which was sufficient for binding. Genetic studies consistent with this result point to an interaction of MED6 (*med6 ts2* mutant allele: Q49L, I68L, L94P, F125Y, R132G, F194L) with MED17 in an area around residue 286 (*srb4-101* allele: E286K) (Lee *et al.*, 1998).

With the same approach of copurification after coexpression the interaction of MED17 and MED8 could be shown for the first time. Coexpression of MED17-F2 with MED8 greatly increases expression and solubility of MED17-F2 (Srb4-F2) and MED8 is purified stoichiometrically along with MED17-F2-His₆ after an affinity chromatography step. Again the residues 241-687 of MED17 (F2 fragment) are sufficient for the interaction.

Bicistronic expression of the MED17-F2 fragment either with MED20 (Srb2), MED19 (Rox3), MED11, or MED22 (Srb6) did not lead to improved expression or solubility. Thus, a large-scale copurification of these complexes was not achieved. Since binding of MED17 to these subunits was predicted by other studies (Kang *et al.*, 2001; Koh *et al.*, 1998; Lee *et al.*, 1998) it is important to emphasize that the

copurification assay is very stringent and selects for specific binding of proteins. Especially if several subunits contribute to an interaction network *in vivo*, lacking of subunits could prevent proper folding and complex assembly in the recombinant system. The MED17 protein and even the MED17-F2 fragment are big polypeptides of 75 and 50 kD respectively. Thus, if MED17 or MED17-F2 is coexpressed with only one binding partner there is still a high probability that large areas of this scaffold protein are unfolded or hydrophobic patches important for interaction are exposed. On the other hand one could speculate on a rather large interface between MED17 and MED6 or MED8 since the bicistronic expression already brings great improvement to solubility.

Another possibility is the binding of head module subunits to the MED17-F1 fragment, which had not been tested here. Since sequences in the F1 fragment are not essential for viability (compare Fig. 2) in yeast (Koh *et al.*, 1998) but comprise stretches that are conserved in evolution it is appealing to speculate that non-essential subunits like MED20 (Srb2) could bind to these sequences. Indeed, binding of MED20 (Srb2) to MED17 (Srb4) has previously been detected in co-immunoprecipitation assays (Kang *et al.*, 2001; Koh *et al.*, 1998; Lee *et al.*, 1998). Thus, interactions between the MED17-F1 fragment and head module subunits would be a good target for further bicistronic expression trials. Using the MED17-F1 fragment would also take advantage of the high expression and solubility of this fragment when expressed from a monocistronic vector. Failure of crystallization of MED17-F1 could be due to partial misfolding or to sequences interfering with crystallization, but might as well simply results from limited trials in finding crystallization conditions. Recent crystallization trials on head module subunits with a crystallization robot point towards a general necessity for high salt concentrations of Mediator head module proteins.

1.2. The MED20/MED18/MED8 proteins form a stable subcomplex within the Mediator head module

A proof of principle for bicistronic expression is provided by comparing MED20 (Srb2) expression levels in combination with MED17-F2 (Srb4) versus MED18 (Srb5). In both cases the gene was cloned untagged in the same vector. Bicistronic expression with a tight interaction partner for MED20 (Srb2), namely MED18 (Srb5),

leads to dramatically increased expression levels and solubility (Fig 6A). While MED20 was expressed but insoluble in combination with MED17-F2 (Fig. 4A), it forms a stoichiometric complex with MED18. The stability of the complex over several purification columns and for weeks at 4°C can be regarded as the best proof for interaction for these two subunits. Conditions for the formation of initial microcrystals hint to high salt concentrations. For crystal improvement additional more specialized screens might already bring about the desired result.

MED18 variants designed according to limited proteolysis results were insoluble. From an alignment with the soluble *S. pombe* MED18 (Srb5) homolog it was inferred, that N-terminal stretches are conserved before a long loop in the MED18 *S. cerevisiae* sequence. Cleavage of the complete N-terminus, as performed according to limited proteolysis results, could thus interfere with proper protein folding. Consistent with proteolytic data a useful strategy might thus be the removal of the internal loop, leaving the conserved N-terminal sequences intact.

It was shown that MED8 binds to the MED20/MED18 subcomplex (Fig. 6c) - a stoichiometric trimeric subcomplex is stable over several copurification steps. Nevertheless, the MED8 protein is not as well expressed and excess of the MED20/MED18 complex needs to be separated from the trimeric complex by gelfiltration. The subcomplex context of MED20/MED18 is not sufficient to assure complete MED8 stability, in a proteolysis assay it is rapidly degraded. These results fit perfectly with the observation that MED8 binds as well to MED17 (Srb4) and that this binding is already sufficient to greatly improve complex solubility. In addition, the interaction between MED8 and MED18 (Srb5) is limited to a very short C-terminal sequence (L. Lariviere, unpublished data).

1.3. Towards reconstitution of the Mediator head module and the Mediator core

Combination of results from bi – and tricistronic expressions led to the postulation and proof of higher order head module subcomplexes. Since MED6 binds to MED17-F2, which in turn binds to MED8 and this to MED20/MED18, a coexpression of two vectors and purification via a GST-tag on MED17-F2 was used to purify a pentameric complex of these subunits (Fig. 10). The purified pentameric subcomplex was very prone to aggregation and was not stoichiometric indicating that hydrophobic patches or unordered sequence stretches, possibly resulting from

missing interaction partners, still interfered with the preparation of a homogenous sample. The presence of the rather big GST-tag on MED17 might also interfere with proper complex assembly. A smaller tag, e.g. His-tag or FLAG tag, possibly on a different subunit (a C-terminal His-tag on MED18 was shown to work in the tricistronic expression), would represent a less invasive strategy. Nevertheless, it was shown that recombinant overexpression of multiple subunits from Mediator head module is possible and the predicted interaction pattern was proved. Since the coexpression of five head module subunits was shown to be possible, the most promising strategy can now be the recombinant expression of the complete head module, thereby circumventing aggregation due to missing subunits. To this end the eight genes coding for head module subunits should be combined on tri- or quatricistronic vectors for coexpression such that known subcomplexes are coded on the same vector.

Most interesting for structural and functional studies on Mediator will be the purification of the core Mediator, comprising the essential subunits. In the presented work it was shown that MED6 binds to MED17 and that this subcomplex can be purified stoichiometrically in larger amounts. These results are consistent with a previously observed functional interaction and a recently published yeast two hybrid interaction map (Guglielmi *et al.*, 2004; Lee and Kim, 1998). MED6, which is an integral part of the head module (Lee and Kim, 1998) and additionally binds directly to the MED7/MED21 subcomplex of the middle module (Baumli *et al.*, 2005) (Fig. 11). This interaction of MED21 with MED6 is apparently essential *in vivo* (Gromoller and Lehming, 2000). 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. Extended loop regions in MED6 indicate a strong intrinsic flexibility that may be crucial for observed large rearrangements within Mediator when it comes to Pol II binding (Asturias *et al.*, 1999; Davis *et al.*, 2002).

Thus, combination of subcomplexes obtained in this study with subcomplexes from the Mediator middle module would make recombinant production of the Mediator core an intermediate term goal. Initial coexpression results with MED21, MED6, MED17 and MED8 (not shown) were already successful, a drawback was again low expression yields and complex aggregation due to missing binding

partners, which can be overcome by stepwise analysis and addition of the minimal combination of subunits for stable purification as shown.

1.4. Structure-function studies on Mediator-RNA Pol II complexes

A future goal in studying the mechanism of Mediator will be structural data of Mediator or individual subunits binding to RNA Pol II. Even though direct binding of Mediator to RNA Pol II has been shown from EM data, it is still an open question which subunits are involved in this interaction. Genetic data from the *srb* genes indicate interaction of these subunits via the CTD, but that could be direct or indirect. A yeast two-hybrid approach for interaction studies between individual subunits of Mediator with individual RNA Pol II subunits was not successful (Guglielmi *et al.*, 2004). A possible combined interaction with between various even small subdomains as well from Mediator as from RNA Pol II can probably not be detected efficiently with this method. Results from EM studies imply that the Mediator head module is the major RNA Pol II interacting submodule of Mediator. The recombinantly purified subcomplexes were thus used in interaction studies with RNA Pol II in order to nail down the interacting subunits, for the use in functional and for obtaining structural data. In preliminary experiments it was shown that the MED17 (Srb4)-MED6 subcomplex coelute with RNA Pol II from a gel filtration column. These data will have to be validated due to limited concentration of MED17-MED6. It was also shown that after an incubation time the trimeric MED8/MED20/MED18 complex coelutes with RNA Pol II from a gel filtration column whereas free trimeric complex elutes in a well separated peak (Fig. 12). The complex is not stoichiometric, indicating weakness of the binding. Whether MED8 is required for binding or whether the non-essential but highly conserved MED20/MED18 complex alone binds to RNA Pol II can now be easily investigated by repeating the binding experiment with the dimeric complex. Although the *med20* (*srb2*) and *med18* (*srb5*) genes are dispensable for cell viability, *in vitro* transcription assays using nuclear extracts from deletion mutant strains reveals that MED20 and MED18 have important roles in basal transcription (Koleske *et al.*, 1992; Thompson *et al.*, 1993; Thompson and Young, 1995), which would be consistent with binding to RNA Pol II. On the other hand MED8 could be responsible for binding RNA Pol II: the C-terminal region was determined to be responsible for Med18 binding (Guglielmi *et al.*, 2004) as well as MED17 binding (Guglielmi *et al.*, 2004). The N-terminal sequences contain large conserved stretches (Boube *et al.*,

2002) that could be responsible for further interactions, namely to MED20 (Guglielmi *et al.*, 2004). The high sensitivity of MED8 in the trimeric complex towards protease digestion observed in this study hints to the importance of the MED8-MED17 interaction. Additional N-terminal sequences could be responsible for RNA Pol II binding. Such questions can be addressed by iterative binding assays of RNA Pol II to head module subcomplexes of different compositions.

The binding of smaller but homogeneously purified subcomplexes to RNA Pol II will additionally be beneficial for structural studies on Mediator-RNA Pol II interaction. On one hand, these samples can be used for EM studies promising higher resolution than the non-homogeneous material that is mostly obtained from native purifications. Varying the combinations of subunits will then lead to a detailed interaction map between Mediator and RNA Pol II. On the other hand use in cocrystallization trials can result in atomic models of the interaction. Using high-resolution structural data of RNA Pol II-Mediator complexes to design appropriate functional studies could bring about the change in understanding the enigmatic mechanism of Mediator function.

2. The Mediator CDK8/Cyclin C module: implications of the Cyclin C structure for function

The CDK8/Cyclin C pair (Srb10/Srb11 in yeast) associates with MED12 (Srb8) and MED13 (Srb9), to form a fourth module of the Mediator that is present in a subpopulation of Mediator complexes. This Mediator module phosphorylates the CTD, is conserved among eukaryotes, and is a target of signal transduction pathways (Borggrefe *et al.*, 2002; Boubé *et al.*, 2002; Liu *et al.*, 2001; Samuelsen *et al.*, 2003). To elucidate the characteristics of this CDK/cyclin pair implied by its association in a tight complex, its different regulation due to no abundance fluctuation by the C-type cyclin and a missing phosphorylation site on CDK8, and its role in activation and repression of transcription, this work presented the structure of Cyclin C from *S. pombe* along with an interaction model with its kinase.

2.1. The Cyclin C structure

In contrast to the subunits of the Mediator head module it was possible to express and purify Cyclin C in a monocistronic expression system. Bicistronic

expression of CDK8 and Cyclin C failed. The kinase was expressed but mostly insoluble, the soluble fraction did apparently not bind to the cyclin, which carried the affinity purification tag. This could be due to missing posttranslational modifications or lack of proper folding in the *E. coli* system. It could also indicate that the MED12 and MED13 subunits stabilize this complex. Therefore an expression and purification protocol for CDK8 with a baculoviral system was established. In order to obtain amounts of CDK8 suitable for structural studies, large-scale expression with this system are the method of choice.

The purified full-length Cyclin C crystallized to small fragile crystals, that were very sensitive to any stress and diffracted only to 3.5 Å. As a first indication of substantial differences in the cyclin fold molecular replacement (AMoRe) with these data using a minimal consensus cyclin model that had been designed after extensive alignments of known cyclin structures failed. Later on, the structure revealed rather high RMSD values especially for the second cyclin repeat (2.5 Å), which might have been the cause. It is important to note that molecular replacement relies on relatively high sequence identity rates as well as good resolution since otherwise noise peaks in the Patterson map interfere with detection of the correct rotation function. Both could not be provided with Cyclin C (*S. pombe*) since it shares only 19% of the sequence with human Cyclin H and since the resolution was very limited. Thus, the calculated potential solutions for the rotation function did not reveal an outstanding best hit. Improved likely-hood enhanced programs that are becoming available might have been able to solve such a problem.

The solution of the Cyclin C structure represents a nice example on how just a few residues might interfere with crystal quality and thus resolution. In contrast to proteolysis assays using trypsin, cleavage with proteinase K pointed to two candidate fragments suitable for crystallization. Whereas removal of the first 29 residues resulted in insoluble protein removing just the first four amino acids drastically improved crystal quality and was the basis for building an atomic model of Cyclin C. Analyzing the obtained protease cleavage sites in the molecular model of Cyclin C reveals the drawbacks as well as the potentials of the limited proteolysis assay: The protease cleaved in a loop connecting the N-terminal helix with the first cyclin repeat. B-factors indicate a high flexibility of this region that probably interferes with crystallization. Removal of such an N-terminal or C-terminal helix would usually be considered as a good candidate for a crystallizable variant. But, probably due to the

highly hydrophobic surface of Cyclin C when expressed without binding partners, the protein uses the helix to dimerize and form a “handshake” interaction thereby shielding hydrophobic patches (Fig. 18B). Cleavage of the helix prevents dimerization and thus leads to insoluble protein, which can be explained by the crystal structure. A similar observation would be expected from cleavage of accessible loops that do not connect independent domains of proteins but rather links to secondary structure elements within the same domain, for example serving as a binding partner platform. This is one of the major reasons why constructs from protease assays may result in insoluble protein (as observed for example with MED18). On the other hand, the N-terminal four amino acids that formed the tip of this “handshake” interaction in the crystal contact made the crystal to fragile due to high flexibility. The assay revealed this cleavage site and removal probably helped to stabilize the crystal contact yielding greatly improved diffraction data.

2.2. Cyclin C differs from known cyclin structures

The exposed position of the N-terminal helix in Cyclin C that is probably a crystallization artifact could nevertheless have implications for function. In contrast to classical cell cycle cyclins like Cyclin A or cyclin E, where the N-terminal helix forms part of the interface with the kinase, a different fold for the N-terminal helix had been observed for the previously solved C-type cyclin structure, namely the structure of Cyclin H. In the case of Cyclin H the N-terminal helix binds tightly to the cyclin repeat. As shown, average B-factors in this area are even lower than in the rest of the protein. Additionally, analysis revealed an interaction highly specific for Cyclin H with residues that are absolutely conserved and at the same time specific for Cyclin H (residues R23, N27, and E53 in human Cyclin H). The authors of the Cyclin H structure state that Cyclin H is not able to activate CDK7 without this helix even though it does not contact the kinase and that this might be due to loss of structural integrity when the protein is mutated. According to B-factor and specificity analysis performed in this work this is likely the case. For the solved cell cycle cyclins it had been observed that the cyclin represent a very rigid structure, where localization of the N-terminal helix is unaltered irrespective of the kinase being bound or not (Card *et al.*, 2000; Jeffrey *et al.*, 1995). With Cyclin C, a third variant of N-terminal helix is presented. Apparently it does not fold towards the interface with the kinase, although this cannot be completely ruled out to be the result of the crystallization artifact. The

fact that no highly conserved residues for the packing of the helix against the C-terminal cyclin repeat were identified together with high B-factors as well as the shortness of the helix point to a different role of the N-terminal helix in Cyclin C. In this context it should be emphasized that the Cyclin C can be regarded as a minimal cyclin. It basically consists only in the cyclin repeats with the additional small N-terminal helix whereas all other cyclins comprise additional N and or C-terminal sequences (compare Fig. 18) of unknown function that are quite long in some cases (e.g. Cyclin A and Cyclin T). It is imaginable that the additional proteins in the Mediator CDK8/Cyclin C module fulfill functions of such domains. Especially the N-terminal helix could adopt a similar fold as observed in the closely related Cyclin H when a binding partner is present. This would also be consistent with an observation that all C-type cyclins possess a conserved W on the N-terminus (Fig. 19), which – in Cyclin H – is involved in the folding. On the other hand, the additional sequences might play a role in the periodical destruction of cyclin via ubiquitination that is responsible for the cycling of cyclins in the cell cycle. Indeed, Cyclin C lacks a destruction box and a PEST sequence, rich in proline (P), glutamate (E), serine (S), and threonine (T) that are present in Cyclin A in these regions (Jacobs *et al.*, 2001).

2.3. A structural model of CDK8 identifies targets for functional analysis

The structure of CDK8 can be modeled due to a high degree of sequence identity among CDKs (37% in the case of CDK2, pdb code 1FIN) and the observation that the backbone fold of all CDK structures solved so far was highly similar (Jeffrey *et al.*, 1995; Lolli *et al.*, 2004; Pavletich, 1999; Tarricone *et al.*, 2001). In combination with an extensive alignment filtering the highly CDK-type specific residues against generally conserved residues, this method allows some predictions on the characteristics of CDK8. Most interesting is the presence of two specific insertions (Fig. 20, 22) the longer one is predicted to be an additional helix. This helix would be located close to the active site. It is easy to imagine the helix being involved in substrate specificity and regulation of CDK8, especially regarding the missing activating phosphorylation site in this kinase class. It remains to be investigated whether the conserved aspartate in the position of the phosphorylation site mimics the phosphorylated state and forms a cluster with the three conserved arginines (Fig. 22C). Modeling showed that the activation segment is 3 residues longer in CDK8 and might thus exhibit a different fold, involving for example the additional helix. A

potential activation of CDK8 in trans by the exposed E91 (Fig.22C) of Cyclin C provides an interesting target for mutational analysis.

Other highly conserved and specific residues cluster in two patches on the CDK8 surface. The first one, with three amino acid residues, is located at the interface with Cyclin C and is probably involved in recognition of it as it was shown by comparison of known CDK/cyclin interfaces (Fig. 21). The second patch is located in an area responsible for inhibitor binding in cell cycle cyclins. Again a clustering of specific residues in this area points to the differences in regulation of this kinase. A similar observation was made in the recently published CDK7 structure (Lolli *et al.*, 2004). The authors state that this area deviates substantially from the inhibitor-binding site in CDK2 and does thus not bind the inhibitor. For CDK7 the region is a good candidate patch for binding MAT1, which is suspected to bind the kinase rather than the cyclin (Andersen *et al.*, 1996b). Binding assays with mutant CDK8 from baculoviral expression and functional studies on substrate specificity after deletion of specific insertions should provide an interesting insight in the function and regulation of the Mediator kinase.

2.4. Analysis of CDK-Cyclin interface suggests the structural basis for cyclin promiscuity

Comparison of the modeled CDK8/Cyclin C interface with known CDK/cyclin interfaces and combination of the results with the data alignment led to the discovery of two parallel patches in the interface: a general recognition patch that could be responsible for nonspecific CDK/cyclin recognition and a patch of specific interactions where a size-complementary interface could be involved in specific recognition of the individual cyclin (Fig. 21). This could explain why, apart from a different time resolution of presence of cyclins, specific CDK/Cyclin Complexes are formed. The C-type cyclins involved in transcription do apparently not fluctuate in abundance throughout the cell cycle. Cyclin A is able to activate CDK7, but this activation is less efficient than with the correct cyclin partner (Andersen *et al.*, 1996a; Lolli *et al.*, 2004). An association of Cyclin C with the human CDK3 during cell cycle (Ren and Rollins, 2004; Sage, 2004) would be imaginable in this context. It should be mentioned that the intrinsic capacity of the cyclin to modulate CDK specificity has recently been emphasized (Loog and Morgan, 2005). In yeast different cyclins that are active throughout the cell cycle associate with the same CDK. This association

itself leads to modulated enzyme activity and could in principle be achieved through alternative conformations in the interface.

2.5. A highly specific groove of Cyclin C

Another mechanism of achieving substrate specificity through the cyclin are substrate recognition patches. Cyclin A possesses a highly specific hydrophobic patch responsible for binding of the RXL motif of substrate proteins (Kontopidis *et al.*, 2003; Schulman *et al.*, 1998). The highly specific groove between the two cyclin repeats discovered in this work could be part of an analogous mechanism for Cyclin C. The different nature of the patch would underline the different requirements for the CTD as substrate, which does probably not contain separated phosphorylation and substrate recognition motifs due to its repetitive nature. Cyclin Ts exhibit a prolonged C-terminal domain that has been shown to be involved in CTD binding via a His-patch (Kurosu *et al.*, 2004; Taube *et al.*, 2002). Alternatively the groove could be involved in establishment of the multiprotein complex this CDK- cyclin pair belongs to, being the binding site for MED12 or MED13. Since CDK8 also phosphorylates other substrates like activators, repressors, TFIID and the Mediator itself (Chi *et al.*, 2001; Nelson *et al.*, 2003; Vincent *et al.*, 2001, Liu *et al.*, 2004, Hallberg *et al.*, 2004) regulation of substrate specificity might – analogous to CDK7 regulation – depend on the macromolecular context. Involvement of the groove in multiprotein complex assembly would explain its specific presence in Cyclin C and absence from Cyclin H. However, it has to be mentioned that a conserved though much smaller patch could also be identified for Cyclin H, located on the C-terminal cyclin repeat and facing the kinase. Several experiments (not shown) were performed to proof binding of the CTD to Cyclin C including GST-pulldown assays with recombinant GST-CTD fusion protein, microcalorimetry with Cyclin C and a two-repeat CTD peptide, and cocrystallization and soaking assays with Cyclin C crystals. Although results were to some extent promising, they could not bring about a definite answer as the obtained signals in all experiments were only slightly above background levels, thus lacking statistical significance. A rather weak binding of the CTD to Cyclin C - microcalorimetry data point to a K_d of around 5 μ M - is on the other hand not surprising for an interaction that has to be transient for function. In cocrystallization experiments some additional density was observed in the Fo-Fc map located close to the specific groove and to a cluster of histidines close to the kinase interface. One

way to improve these results and increase the potentially low occupancy in the crystal resulting from weak interaction could be a direct fusion of a CTD repeat to the Cyclin C molecule. Additional density resulting from a well-defined interaction would then strongly support a hypothesis of Cyclin C binding the CTD rather than other substrates or Mediator subunits.

2.6. The CDK8/Cyclin C module and its role in Mediator function

The CDK8/Cyclin C module is variably present in different yeast Mediator preparations (Borggreve *et al.*, 2002; Samuelsen *et al.*, 2003), RNA Pol II was shown to interact only with Mediator lacking the CDK8/Cyclin C module (Samuelsen *et al.*, 2003) and similar observations have been made for mammalian Mediator (Naar *et al.*, 2002; Wang *et al.*, 2001). Since Mediator containing the module represses basal transcription in vitro (Spahr *et al.*, 2003; Taatjes *et al.*, 2002) whereas Mediator lacking it has a stimulatory effect (Kim *et al.*, 1994), presence of the CDK8/Cyclin C module might well provide a switch from repression to activation during transcription initiation (Bjorklund and Gustafsson, 2005). It was recently suggested that corresponding forms of Mediator might be recruited by activators and repressors to regulatory sequences rather than to the promoter itself, whereas RNA Pol II and GTFs are recruited to the promoter 2-3 min later (Bryant and Ptashne, 2003) contradicting the proposition of a preformed holoenzyme complex (Bjorklund and Gustafsson, 2005; Malik and Roeder, 2005).

Conformational changes within Mediator that have been observed, and for which structural evidence has been presented (Baumli *et al.*, 2005), might trigger a mechanism in which CDK8/Cyclin C containing Mediator is unable to bind RNA Pol II and thereby inhibits transcription. If dissociation of the CDK8/Cyclin C module was to confer conformational changes in the Mediator it should bind in the vicinity of subunits that have so far been implied in this movement. Indeed, contacts between kinase and middle module (MED1 and MED4) have been shown (Kang *et al.*, 2001) and an interaction between Srb9 and MED17 (Srb4) has recently been described (Guglielmi *et al.*, 2004). In addition, the tail module serves as a substrate for CDK8, MED2 is a phosphorylation target (Hallberg *et al.*, 2004).

When electron microscopy data of Mediator with and without the CDK8/Cyclin C module become available it will be very interesting to determine the position and thus maybe function of the CDK- cyclin pair by docking the crystal structure and the

model into the density. This should also answer the questions whether the CDK8/Cyclin C module is involved in conformational changes of Mediator.

References

- Adamczewski, J. P., Rossignol, M., Tassan, J. P., Nigg, E. A., Moncollin, V., and Egly, J. M. (1996). MAT1, cdk7 and cyclin H form a kinase complex which is UV light-sensitive upon association with TFIIH. *Embo J* 15, 1877-1884.
- Akoulitchev, S., Chuikov, S., and Reinberg, D. (2000). TFIIH is negatively regulated by cdk8-containing mediator complexes. *Nature* 407, 102-106.
- Albright, S. R., and Tjian, R. (2000). TAFs revisited: more data reveal new twists and confirm old ideas. *Gene* 242, 1-13.
- Andersen, G., Busso, D., Poterszman, A., Hwang, J. R., Wurtz, J.-M., Ripp, R., Thierry, J.-C., Egly, J.-M., and Moras, D. (1996a). The structure of cyclin H: Common mode of kinase activation and specific features. *EMBO*.
- 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, 14706-14709.
- Armache, K. J., Mitterweger, S., Meinhart, A., and Cramer, P. (2005). Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. *J Biol Chem* 280, 7131-7134.
- 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, 985-987.
- Barton, G. J. (1993). ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng* 6, 37-40.
- Baumli, S., Hoepfner, S., and Cramer, P. (2005). A conserved mediator hinge revealed in the structure of the MED7.MED21 (Med7.Srb7) heterodimer. *J Biol Chem* 280, 18171-18178.
- Bensaude, O., Bonnet, F., Casse, C., Dubois, M. F., Nguyen, V. T., and Palancade, B. (1999). Regulated phosphorylation of the RNA polymerase II C-terminal domain (CTD). *Biochem Cell Biol* 77, 249-255.
- Bentley, D. (2002). The mRNA assembly line: transcription and processing machines in the same factory. *Curr Opin Cell Biol* 14, 336-342.
- Bjorklund, S., and Gustafsson, C. M. (2004). The mediator complex. *Adv Protein Chem* 67, 43-65.
- Bjorklund, S., and Gustafsson, C. M. (2005). The yeast Mediator complex and its regulation. *Trends Biochem Sci* 30, 240-244.
- Bloom H., Beier H., Gross H.J. (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, 44202-44207.
- 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, 143-151.
- 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., *et al.* (2004). A unified nomenclature for protein subunits of mediator complexes linking transcriptional regulators to RNA polymerase II. *Mol Cell* 14, 553-557.
- Bregman, D. B., Pestell, R. G., and Kidd, V. J. (2000). Cell cycle regulation and RNA polymerase II. *Front Biosci* 5, D244-257.
- Brodolin, K., Zenkin, N., Mustaev, A., Mamaeva, D., and Heumann, H. (2004). The sigma 70 subunit of RNA polymerase induces lacUV5 promoter-proximal pausing of transcription. *Nat Struct Mol Biol* 11, 551-557.
- 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., *et al.* (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 54, 905-921.
- Bryant, G. O., and Ptashne, M. (2003). Independent recruitment in vivo by Gal4 of two complexes required for transcription. *Mol Cell* 11, 1301-1309.
- Budisa, N., Steipe, B., Demange, P., Eckerskorn, C., Kellermann, J., and Huber, R. (1995). High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in *Escherichia coli*. *Eur J Biochem* 230, 788-796.
- Buratowski, S. (1994). The basics of basal transcription by RNA polymerase II. *Cell* 77, 1-3.
- Cadena, D. L., and Dahmus, M. E. (1987). Messenger RNA synthesis in mammalian cells is catalyzed by the phosphorylated form of RNA polymerase II. *J Biol Chem* 262, 12468-12474.
- Card, G. L., Knowles, P., Laman, H., Jones, N., and McDonald, N. Q. (2000). Crystal structure of a gamma-herpesvirus cyclin-cdk complex. *Embo J* 19, 2877-2888.
- CCP4 (1994). The CCP4 Suite: Programs for protein crystallography. *Acta Cryst D* 50, 760-763.
- Chang, Y. W., Howard, S. C., and Herman, P. K. (2004). The Ras/PKA signaling pathway directly targets the Srb9 protein, a component of the general RNA polymerase II transcription apparatus. *Mol Cell* 15, 107-116.

- Chi, Y., Huddleston, M. J., Zhang, X., Young, R. A., Annan, R. S., Carr, S. A., and Deshaies, R. J. (2001). Negative regulation of Gcn4 and Msn2 transcription factors by Srb10 cyclin-dependent kinase. *Genes Dev* *15*, 1078-1092.
- Cho, E. J., Kobor, M. S., Kim, M., Greenblatt, J., and Buratowski, S. (2001). Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. *Genes Dev* *15*, 3319-3329.
- Cho, E. J., Takagi, T., Moore, C. R., and Buratowski, S. (1997). mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev* *11*, 3319-3326.
- Conaway, R. C., and Conaway, J. W. (1993). General initiation factors for RNA polymerase II. *Annu Rev Biochem* *62*, 161-190.
- Corden, J. L., Cadena, D. L., Ahearn, J. M., Jr., and Dahmus, M. E. (1985). A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. *Proc Natl Acad Sci U S A* *82*, 7934-7938.
- Cosma, M. P., Tanaka, T., and Nasmyth, K. (1999). Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter. *Cell* *97*, 299-311.
- Cramer, P., Bushnell, D. A., and Kornberg, R. D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* *292*, 1863-1876.
- 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*, 409-415.
- Devault, A., Martinez, A. M., Fesquet, D., Labbe, J. C., Morin, N., Tassan, J. P., Nigg, E. A., Cavadore, J. C., and Doree, M. (1995). MAT1 ('menage a trois') a new RING finger protein subunit stabilizing cyclin H-cdk7 complexes in starfish and *Xenopus* CAK. *Embo J* *14*, 5027-5036.
- 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*, 14307-14310.
- Dubois, M. F., Vincent, M., Vigneron, M., Adamczewski, J., Egly, J. M., and Bensaude, O. (1997). Heat-shock inactivation of the TFIIH-associated kinase and change in the phosphorylation sites on the C-terminal domain of RNA polymerase II. *Nucleic Acids Res* *25*, 694-700.
- Dynlacht, B. D. (1997). Regulation of transcription by proteins that control the cell cycle. *Nature* *389*, 149-152.
- Ebright, R. H. (2000). RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. *J Mol Biol* *304*, 687-698.

- Edwards, M. C., Wong, C., and Elledge, S. J. (1998). Human cyclin K, a novel RNA polymerase II-associated cyclin possessing both carboxy-terminal domain kinase and Cdk-activating kinase activity. *Mol Cell Biol* 18, 4291-4300.
- Fisher, R. P., Jin, P., Chamberlin, H. M., and Morgan, D. O. (1995). Alternative mechanisms of CAK assembly require an assembly factor or an activating kinase. *Cell* 83, 47-57.
- Flanagan, P. M., Kelleher, R. J., 3rd, Sayre, M. H., Tschochner, H., and Kornberg, R. D. (1991). A mediator required for activation of RNA polymerase II transcription in vitro. *Nature* 350, 436-438.
- 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, 8329-8333.
- Fondell, J. D., Guermah, M., Malik, S., and Roeder, R. G. (1999). Thyroid hormone receptor-associated proteins and general positive cofactors mediate thyroid hormone receptor function in the absence of the TATA box-binding protein-associated factors of TFIID. *Proc Natl Acad Sci U S A* 96, 1959-1964.
- Fryer, C. J., White, J. B., and Jones, K. A. (2004). Mastermind recruits CycC:CDK8 to phosphorylate the Notch ICD and coordinate activation with turnover. *Mol Cell* 16, 509-520.
- Fu, T. J., Peng, J., Lee, G., Price, D. H., and Flores, O. (1999). Cyclin K functions as a CDK9 regulatory subunit and participates in RNA polymerase II transcription. *J Biol Chem* 274, 34527-34530.
- Garriga, J., Peng, J., Parreno, M., Price, D. H., Henderson, E. E., and Grana, X. (1998). Upregulation of cyclin T1/CDK9 complexes during T cell activation. *Oncogene* 17, 3093-3102.
- Gill, G., and Ptashne, M. (1988). Negative effect of the transcriptional activator GAL4. *Nature* 334, 721-724.
- Gromoller, A., and Lehming, N. (2000). Srb7p is essential for the activation of a subset of genes. *FEBS Lett* 484, 48-54.
- 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, 97-108.
- 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, 5379- 5391.
- Guo, Z., and Stiller, J. W. (2004). Comparative genomics of cyclin-dependent kinases suggest co-evolution of the RNAP II C-terminal domain and CTD-directed CDKs. *BMC Genomics* 5, 69.

- Gustafsson, C. M., and Samuelsson, T. (2001). Mediator--a universal complex in transcriptional regulation. *Mol Microbiol* *41*, 1-8.
- Hahn, S. (2004). Structure and mechanism of the RNA polymerase II transcription machinery. *Nat Struct Mol Biol* *11*, 394-403.
- 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*, 3370-3375.
- 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*, 979-988.
- Happel, A. M., Swanson, M. S., and Winston, F. (1991). The SNF2, SNF5 and SNF6 genes are required for Ty transcription in *Saccharomyces cerevisiae*. *Genetics* *128*, 69-77.
- Harper, J. W., and Elledge, S. J. (1998). The role of Cdk7 in CAK function, a retro-retrospective. *Genes Dev* *12*, 285-289.
- 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*, 43-53.
- Hermant, D., Pihlak, A., Westerling, T., Damagnez, V., Vandenhoute, J., Cottarel, G., and Makela, T. P. (1998). Fission yeast Csk1 is a CAK-activating kinase (CAKAK). *Embo J* *17*, 7230-7238.
- Higuchi, R., Krummel, B., and Saiki, R. K. (1988). A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Res* *16*, 7351-7367.
- Ho, C. K., Sriskanda, V., McCracken, S., Bentley, D., Schwer, B., and Shuman, S. (1998). The guanylyltransferase domain of mammalian mRNA capping enzyme binds to the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J Biol Chem* *273*, 9577-9585.
- 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*, 717-728.
- Honda, R., Lowe, E. D., Dubinina, E., Skamnaki, V., Cook, A., Brown, N. R., and Johnson, L. N. (2005). The structure of cyclin E1/CDK2: implications for CDK2 activation and CDK2-independent roles. *Embo J* *24*, 452-463.
- Hu, D., Mayeda, A., Trembley, J. H., Lahti, J. M., and Kidd, V. J. (2003). CDK11 complexes promote pre-mRNA splicing. *J Biol Chem* *278*, 8623-8629.

- 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, 4569-4574.
- Jacobs, H. W., Keidel, E., and Lehner, C. F. (2001). A complex degradation signal in Cyclin A required for G1 arrest, and a C-terminal region for mitosis. *Embo J* 20, 2376-2386.
- Jeffrey, P. D., Russo, A. A., Polyak, K., Gibbs, E., Hurwitz, J., Massague, J., and Pavletich, N. P. (1995). Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature* 376, 313-320.
- 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, 8538-8543.
- Jones, J. C., Phatnani, H. P., Haystead, T. A., MacDonald, J. A., Alam, S. M., and Greenleaf, A. L. (2004). C-terminal repeat domain kinase I phosphorylates Ser2 and Ser5 of RNA polymerase II C-terminal domain repeats. *J Biol Chem* 279, 24957-24964.
- 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), 110-119.
- Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22, 2577-2637.
- Kaldis, P., Sutton, A., and Solomon, M. J. (1996). The Cdk-activating kinase (CAK) from budding yeast. *Cell* 86, 553-564.
- 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, 42003-42010.
- Kang, M. E., and Dahmus, M. E. (1993). RNA polymerases IIA and IIO have distinct roles during transcription from the TATA-less murine dihydrofolate reductase promoter. *J Biol Chem* 268, 25033-25040.
- Kelleher, R. J., 3rd, Flanagan, P. M., and Kornberg, R. D. (1990). A novel mediator between activator proteins and the RNA polymerase II transcription apparatus. *Cell* 61, 1209-1215.
- Keogh, M. C., Podolny, V., and Buratowski, S. (2003). Bur1 kinase is required for efficient transcription elongation by RNA polymerase II. *Mol Cell Biol* 23, 7005-7018.

- Kettenberger, H., Armache, K. J., and Cramer, P. (2004). Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. *Mol Cell* 16, 955-965.
- Kim, Y. J., Bjorklund, S., Li, Y., Sayre, M. H., and Kornberg, R. D. (1994). A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* 77, 599-608.
- Kim, Y. K., Bourgeois, C. F., Isel, C., Churcher, M. J., and Karn, J. (2002). Phosphorylation of the RNA polymerase II carboxyl-terminal domain by CDK9 is directly responsible for human immunodeficiency virus type 1 Tat-activated transcriptional elongation. *Mol Cell Biol* 22, 4622-4637.
- 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, 895-904.
- Koleske, A. J., Buratowski, S., Nonet, M., and Young, R. A. (1992). A novel transcription factor reveals a functional link between the RNA polymerase II CTD and TFIID. *Cell* 69, 883-894.
- Koleske, A. J., and Young, R. A. (1994). An RNA polymerase II holoenzyme responsive to activators. *Nature* 368, 466-469.
- Koleske, A. J., and Young, R. A. (1995). The RNA polymerase II holoenzyme and its implications for gene regulation. *Trends Biochem Sci* 20, 113-116.
- Komarnitsky, P., Cho, E. J., and Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev* 14, 2452-2460.
- Kontopidis, G., Andrews, M. J., McInnes, C., Cowan, A., Powers, H., Innes, L., Plater, A., Griffiths, G., Paterson, D., Zheleva, D. I., *et al.* (2003). Insights into cyclin groove recognition: complex crystal structures and inhibitor design through ligand exchange. *Structure (Camb)* 11, 1537-1546.
- Kornberg, R. D. (2005). Mediator and the mechanism of transcriptional activation. *Trends Biochem Sci* 30, 235-239.
- 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, 3927-3937.
- Kurosu, T., Zhang, F., and Peterlin, B. M. (2004). Transcriptional activity and substrate recognition of cyclin T2 from P-TEFb. *Gene* 343, 173-179.
- La Fortelle, E. d., 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, 680-685.

- Larochelle, S., Pandur, J., Fisher, R. P., Salz, H. K., and Suter, B. (1998). Cdk7 is essential for mitosis and for in vivo Cdk-activating kinase activity. *Genes Dev* 12, 370-381.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993a). PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst* 26, 283-291.
- Laskowski, R. A., Moss, D. S., and Thornton, J. M. (1993b). Main-chain bond lengths and bond angles in protein structures. *J Mol Biol* 231, 1049-1067.
- Lee, J. M., and Greenleaf, A. L. (1991). CTD kinase large subunit is encoded by CTK1, a gene required for normal growth of *Saccharomyces cerevisiae*. *Gene Expr* 1, 149-167.
- Lee, K. M., Saiz, J. E., Barton, W. A., and Fisher, R. P. (1999a). Cdc2 activation in fission yeast depends on Mcs6 and Csk1, two partially redundant Cdk-activating kinases (CAKs). *Curr Biol* 9, 441-444.
- Lee, T. I., Wyrick, J. J., Koh, S. S., Jennings, E. G., Gadbois, E. L., and Young, R. A. (1998). Interplay of positive and negative regulators in transcription initiation by RNA polymerase II holoenzyme. *Mol Cell Biol* 18, 4455-4462.
- 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, 5364-5370.
- Lee, Y. C., Park, J. M., Min, S., Han, S. J., and Kim, Y. J. (1999b). An activator binding module of yeast RNA polymerase II holoenzyme. *Mol Cell Biol* 19, 2967-2976.
- Lees, E. (1995). Cyclin dependent kinase regulation. *Curr Opin Cell Biol* 7, 773-780.
- Leopold, P., and O'Farrell, P. H. (1991). An evolutionarily conserved cyclin homolog from *Drosophila* rescues yeast deficient in G1 cyclins. *Cell* 66, 1207-1216.
- Lew, D. J., Dulic, V., and Reed, S. I. (1991). Isolation of three novel human cyclins by rescue of G1 cyclin (Cln) function in yeast. *Cell* 66, 1197-1206.
- 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, 10864-10868.
- Liao, S. M., Zhang, J., Jeffery, D. A., Koleske, A. J., Thompson, C. M., Chao, D. M., Viljoen, M., van Vuuren, H. J., and Young, R. A. (1995). A kinase-cyclin pair in the RNA polymerase II holoenzyme. *Nature* 374, 193-196.
- 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, 49455-49459.

- 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, 1721-1735.
- 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, 7169-7175.
- Lolli, G., Lowe, E. D., Brown, N. R., and Johnson, L. N. (2004). The crystal structure of human CDK7 and its protein recognition properties. *Structure (Camb)* 12, 2067-2079.
- Loog, M., and Morgan, D. O. (2005). Cyclin specificity in the phosphorylation of cyclin-dependent kinase substrates. *Nature* 434, 104-108.
- Lorch, Y., Beve, J., Gustafsson, C. M., Myers, L. C., and Kornberg, R. D. (2000). Mediator-nucleosome interaction. *Mol Cell* 6, 197-201.
- Lu, H., Zawel, L., Fisher, L., Egly, J. M., and Reinberg, D. (1992). Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature* 358, 641-645.
- 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, 387-397.
- Malik, S., and Roeder, R. G. (2000). Transcriptional regulation through Mediator-like coactivators in yeast and metazoan cells. *Trends Biochem Sci* 25, 277-283.
- Malik, S., and Roeder, R. G. (2005). Dynamic regulation of pol II transcription by the mammalian Mediator complex. *Trends Biochem Sci* 30, 256-263.
- Marshall, N. F., Peng, J., Xie, Z., and Price, D. H. (1996). Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. *J Biol Chem* 271, 27176-27183.
- Marshall, N. F., and Price, D. H. (1995). Purification of P-TEFb, a transcription factor required for the transition into productive elongation. *J Biol Chem* 270, 12335-12338.
- McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S., and Bentley, D. L. (1997). 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev* 11, 3306-3318.
- Meinhart, A., Blobel, J., and Cramer, P. (2003). An extended winged helix domain in general transcription factor E/IE alpha. *J Biol Chem* 278, 48267-48274.
- Meinhart, A., and Cramer, P. (2004). Recognition of RNA polymerase II carboxy-terminal domain by 3'-RNA-processing factors. *Nature* 430, 223-226.

- Meinhart, A., Kamenski, T., Hoepfner, S., Baumli, S., and Cramer, P. (2005). A structural perspective of CTD function. *Genes Dev* 19, 1401-1415.
- 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, 981-993.
- Michels, A. A., Nguyen, V. T., Fraldi, A., Labas, V., Edwards, M., Bonnet, F., Lania, L., and Bensaude, O. (2003). MAQ1 and 7SK RNA interact with CDK9/cyclin T complexes in a transcription-dependent manner. *Mol Cell Biol* 23, 4859-4869.
- Molz, L., and Beach, D. (1993). Characterization of the fission yeast *mcs2* cyclin and its associated protein kinase activity. *Embo J* 12, 1723-1732.
- Morgan, D. O. (1995). Principles of CDK regulation. *Nature* 374, 131-134.
- Mukhopadhyay, J., Kapanidis, A. N., Mekler, V., Kortkhonjia, E., Ebright, Y. W., and Ebright, R. H. (2001). Translocation of sigma(70) with RNA polymerase during transcription: fluorescence resonance energy transfer assay for movement relative to DNA. *Cell* 106, 453-463.
- Murray, A. W. (2004). Recycling the cell cycle: cyclins revisited. *Cell* 116, 221-234.
- Murray, S., Udupa, R., Yao, S., Hartzog, G., and Prelich, G. (2001). Phosphorylation of the RNA polymerase II carboxy-terminal domain by the Bur1 cyclin-dependent kinase. *Mol Cell Biol* 21, 4089-4096.
- Myer, V. E., and Young, R. A. (1998). RNA polymerase II holoenzymes and subcomplexes. *J Biol Chem* 273, 27757-27760.
- 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, 45-54.
- Myers, L. C., and Kornberg, R. D. (2000). Mediator of transcriptional regulation. *Annu Rev Biochem* 69, 729-749.
- Naar, A. M., Beurang, P. A., Zhou, S., Abraham, S., Solomon, W., and Tjian, R. (1999). Composite co-activator ARC mediates chromatin-directed transcriptional activation. *Nature* 398, 828-832.
- 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, 1339-1344.
- Neigeborn, L., Celenza, J. L., and Carlson, M. (1987). SSN20 is an essential gene with mutant alleles that suppress defects in SUC2 transcription in *Saccharomyces cerevisiae*. *Mol Cell Biol* 7, 672-678.

- Nelson, C., Goto, S., Lund, K., Hung, W., and Sadowski, I. (2003). Srb10/Cdk8 regulates yeast filamentous growth by phosphorylating the transcription factor Ste12. *Nature* 421, 187-190.
- Nguyen, V. T., Kiss, T., Michels, A. A., and Bensaude, O. (2001). 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature* 414, 322-325.
- Nigg, E. A. (1996). Cyclin-dependent kinase 7: at the cross-roads of transcription, DNA repair and cell cycle control? *Curr Opin Cell Biol* 8, 312-317.
- Nikolov, D. B., Chen, H., Halay, E. D., Usheva, A. A., Hisatake, K., Lee, D. K., Roeder, R. G., and Burley, S. K. (1995). Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* 377, 119-128.
- Nishiwaki, E., Turner, S. L., Harju, S., Miyazaki, S., Kashiwagi, M., Koh, J., and Serizawa, H. (2000). Regulation of CDK7-carboxyl-terminal domain kinase activity by the tumor suppressor p16(INK4A) contributes to cell cycle regulation. *Mol Cell Biol* 20, 7726-7734.
- Nonet, M., Sweetser, D., and Young, R. A. (1987). Functional redundancy and structural polymorphism in the large subunit of RNA polymerase II. *Cell* 50, 909-915.
- 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, 715-724.
- Oelgeschlager, T. (2002). Regulation of RNA polymerase II activity by CTD phosphorylation and cell cycle control. *J Cell Physiol* 190, 160-169.
- Orphanides, G., Lagrange, T., and Reinberg, D. (1996). The general transcription factors of RNA polymerase II. *Genes Dev* 10, 2657-2683.
- Otwinowski, Z., and Minor, W. (1996). Processing of X-ray diffraction data collected in oscillation mode. *Meth Enzym* 276, 307-326.
- Park, J. M., Gim, B. S., Kim, J. M., Yoon, J. H., Kim, H. S., Kang, J. G., and Kim, Y. J. (2001). *Drosophila* Mediator complex is broadly utilized by diverse gene-specific transcription factors at different types of core promoters. *Mol Cell Biol* 21, 2312-2323.
- 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, 8709-8719.
- Patturajan, M., Schulte, R. J., Sefton, B. M., Berezney, R., Vincent, M., Bensaude, O., Warren, S. L., and Corden, J. L. (1998). Growth-related changes in phosphorylation of yeast RNA polymerase II. *J Biol Chem* 273, 4689-4694.

- Pavletich, N. P. (1999). Mechanisms of cyclin-dependent kinase regulation: structures of Cdks, their cyclin activators, and Cip and INK4 inhibitors. *J Mol Biol* 287, 821-828.
- Payne, J. M., and Dahmus, M. E. (1993). Partial purification and characterization of two distinct protein kinases that differentially phosphorylate the carboxyl-terminal domain of RNA polymerase subunit IIa. *J Biol Chem* 268, 80-87.
- Peng, J., Zhu, Y., Milton, J. T., and Price, D. H. (1998). Identification of multiple cyclin subunits of human P-TEFb. *Genes Dev* 12, 755-762.
- 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, 1004-1014.
- Powell, H. R. (1999). The Rossmann Fourier autoindexing algorithm in MOSFLM. *Acta Crystallogr D Biol Crystallogr* 55 (Pt 10), 1690-1695.
- Prelich, G. (2002). RNA polymerase II carboxy-terminal domain kinases: emerging clues to their function. *Eukaryot Cell* 1, 153-162.
- Prelich, G., and Winston, F. (1993). Mutations that suppress the deletion of an upstream activating sequence in yeast: involvement of a protein kinase and histone H3 in repressing transcription in vivo. *Genetics* 135, 665-676.
- Rachez, C., and Freedman, L. P. (2001). Mediator complexes and transcription. *Curr Opin Cell Biol* 13, 274-280.
- 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, 824-828.
- Ramanathan, Y., Rajpara, S. M., Reza, S. M., Lees, E., Shuman, S., Mathews, M. B., and Pe'ery, T. (2001). Three RNA polymerase II carboxyl-terminal domain kinases display distinct substrate preferences. *J Biol Chem* 276, 10913-10920.
- Ramanathan, Y., Reza, S. M., Young, T. M., Mathews, M. B., and Pe'ery, T. (1999). Human and rodent transcription elongation factor P-TEFb: interactions with human immunodeficiency virus type 1 tat and carboxy-terminal domain substrate. *J Virol* 73, 5448-5458.
- Rani, P. G., Ranish, J. A., and Hahn, S. (2004). RNA polymerase II (Pol II)-TFIIF and Pol II-mediator complexes: the major stable Pol II complexes and their activity in transcription initiation and reinitiation. *Mol Cell Biol* 24, 1709-1720.
- 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, 49-63.

- Ren, S., and Rollins, B. J. (2004). Cyclin C/cdk3 promotes Rb-dependent G0 exit. *Cell* 117, 239-251.
- Rickert, P., Corden, J. L., and Lees, E. (1999). Cyclin C/CDK8 and cyclin H/CDK7/p36 are biochemically distinct CTD kinases. *Oncogene* 18, 1093-1102.
- Rickert, P., Seghezzi, W., Shanahan, F., Cho, H., and Lees, E. (1996). Cyclin C/CDK8 is a novel CTD kinase associated with RNA polymerase II. *Oncogene* 12, 2631-2640.
- Roeder, R. G. (1996). The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem Sci* 21, 327-335.
- Roeder, R. G. (1998). Role of general and gene-specific cofactors in the regulation of eukaryotic transcription. *Cold Spring Harb Symp Quant Biol* 63, 201-218.
- Rossignol, M., Kolb-Cheynel, I., and Egly, J. M. (1997). Substrate specificity of the cdk-activating kinase (CAK) is altered upon association with TFIIH. *Embo J* 16, 1628-1637.
- Russo, A. A., Jeffrey, P. D., and Pavletich, N. P. (1996). Structural basis of cyclin-dependent kinase activation by phosphorylation. *Nat Struct Biol* 3, 696-700.
- Russo, A. A., Tong, L., Lee, J. O., Jeffrey, P. D., and Pavletich, N. P. (1998). Structural basis for inhibition of the cyclin-dependent kinase Cdk6 by the tumour suppressor p16INK4a. *Nature* 395, 237-243.
- Ryu, S., and Tjian, R. (1999). Purification of transcription cofactor complex CRSP. *Proc Natl Acad Sci U S A* 96, 7137-7142.
- Ryu, S., Zhou, S., Ladurner, A. G., and Tjian, R. (1999). The transcriptional cofactor complex CRSP is required for activity of the enhancer-binding protein Sp1. *Nature* 397, 446-450.
- Sage, J. (2004). Cyclin C makes an entry into the cell cycle. *Dev Cell* 6, 607-608.
- Sambrook J., Russel D.W. (2001). *Molecular Cloning: a laboratory manual*, 3rd edison. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press.
- Samuelson, 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, 6422-6427.
- Sato, S., Tomomori-Sato, C., Banks, C. A., Sorokina, I., Parmely, T. J., Kong, S. E., Jin, J., Cai, Y., Lane, W. S., Brower, C. S., *et al.* (2003). Identification of mammalian Mediator subunits with similarities to yeast Mediator subunits Srb5, Srb6, Med11, and Rox3. *J Biol Chem* 278, 15123-15127.

- Schroeder, S. C., Schwer, B., Shuman, S., and Bentley, D. (2000). Dynamic association of capping enzymes with transcribing RNA polymerase II. *Genes Dev* 14, 2435-2440.
- Schulman, B. A., Lindstrom, D. L., and Harlow, E. (1998). Substrate recruitment to cyclin-dependent kinase 2 by a multipurpose docking site on cyclin A. *Proc Natl Acad Sci U S A* 95, 10453-10458.
- Schulze-Gahmen, U., and Kim, S. H. (2002). Structural basis for CDK6 activation by a virus-encoded cyclin. *Nat Struct Biol* 9, 177-181.
- Shilatifard, A., Conaway, R. C., and Conaway, J. W. (2003). The RNA polymerase II elongation complex. *Annu Rev Biochem* 72, 693-715.
- Shim, E. Y., Walker, A. K., Shi, Y., and Blackwell, T. K. (2002). CDK-9/cyclin T (P-TEFb) is required in two postinitiation pathways for transcription in the *C. elegans* embryo. *Genes Dev* 16, 2135-2146.
- 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, 1351-1356.
- Spahr, H., Khorosjutina, O., Baraznenok, V., Linder, T., Samuelson, 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, 51301-51306.
- Spahr, H., Samuelson, C. O., Baraznenok, V., Ernest, I., Huylebroeck, D., Remacle, J. E., Samuelsson, T., Kieselbach, T., Holmberg, S., and Gustafsson, C. M. (2001). Analysis of *Schizosaccharomyces pombe* mediator reveals a set of essential subunits conserved between yeast and metazoan cells. *Proc Natl Acad Sci U S A* 98, 11985-11990.
- Sterner, D. E., Lee, J. M., Hardin, S. E., and Greenleaf, A. L. (1995). The yeast carboxyl-terminal repeat domain kinase CTDK-I is a divergent cyclin-cyclin-dependent kinase complex. *Mol Cell Biol* 15, 5716-5724.
- Storoni, L. C., McCoy, A. J., and Read, R. J. (2004). Likelihood-enhanced fast rotation functions. *Acta Crystallogr D Biol Crystallogr* 60, 432-438.
- 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, 213-222.
- Svejstrup, J. Q., Li, Y., Fellows, J., Gnatt, A., Bjorklund, S., and Kornberg, R. D. (1997). Evidence for a mediator cycle at the initiation of transcription. *Proc Natl Acad Sci U S A* 94, 6075-6078.
- Swanson, M. S., Malone, E. A., and Winston, F. (1991). SPT5, an essential gene important for normal transcription in *Saccharomyces cerevisiae*, encodes an acidic nuclear protein with a carboxy-terminal repeat. *Mol Cell Biol* 11, 4286.

- Swanson, M. S., and Winston, F. (1992). SPT4, SPT5 and SPT6 interactions: effects on transcription and viability in *Saccharomyces cerevisiae*. *Genetics* 132, 325-336.
- Taatjes, D. J., Naar, A. M., Andel, F., 3rd, Nogales, E., and Tjian, R. (2002). Structure, function, and activator-induced conformations of the CRSP coactivator. *Science* 295, 1058-1062.
- Tarricone, C., Dhavan, R., Peng, J., Areces, L. B., Tsai, L. H., and Musacchio, A. (2001). Structure and regulation of the CDK5-p25(nck5a) complex. *Mol Cell* 8, 657-669.
- Tassan, J. P., Jaquenoud, M., Leopold, P., Schultz, S. J., and Nigg, E. A. (1995). Identification of human cyclin-dependent kinase 8, a putative protein kinase partner for cyclin C. *Proc Natl Acad Sci U S A* 92, 8871-8875.
- Tassan, J. P., Schultz, S. J., Bartek, J., and Nigg, E. A. (1994). Cell cycle analysis of the activity, subcellular localization, and subunit composition of human CAK (CDK-activating kinase). *J Cell Biol* 127, 467-478.
- Taube, R., Lin, X., Irwin, D., Fujinaga, K., and Peterlin, B. M. (2002). Interaction between P-TEFb and the C-terminal domain of RNA polymerase II activates transcriptional elongation from sites upstream or downstream of target genes. *Mol Cell Biol* 22, 321-331.
- Terwilliger, T. C. (2002). Automated structure solution, density modification and model building. *Acta Crystallogr D Biol Crystallogr* 58, 1937-1940.
- Thompson, C. M., Koleske, A. J., Chao, D. M., and Young, R. A. (1993). A multisubunit complex associated with the RNA polymerase II CTD and TATA-binding protein in yeast. *Cell* 73, 1361-1375.
- Thompson, C. M., and Young, R. A. (1995). General requirement for RNA polymerase II holoenzymes in vivo. *Proc Natl Acad Sci U S A* 92, 4587-4590.
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22, 4673-4680.
- Tomomori-Sato, C., Sato, S., Parmely, T. J., Banks, C. A., Sorokina, I., Florens, L., Zybailov, B., Washburn, M. P., Brower, C. S., Conaway, R. C., and Conaway, J. W. (2004). A mammalian mediator subunit that shares properties with *Saccharomyces cerevisiae* mediator subunit Cse2. *J Biol Chem* 279, 5846-5851.
- Trembley, J. H., Hu, D., Slaughter, C. A., Lahti, J. M., and Kidd, V. J. (2003). Casein kinase 2 interacts with cyclin-dependent kinase 11 (CDK11) in vivo and phosphorylates both the RNA polymerase II carboxyl-terminal domain and CDK11 in vitro. *J Biol Chem* 278, 2265-2270.

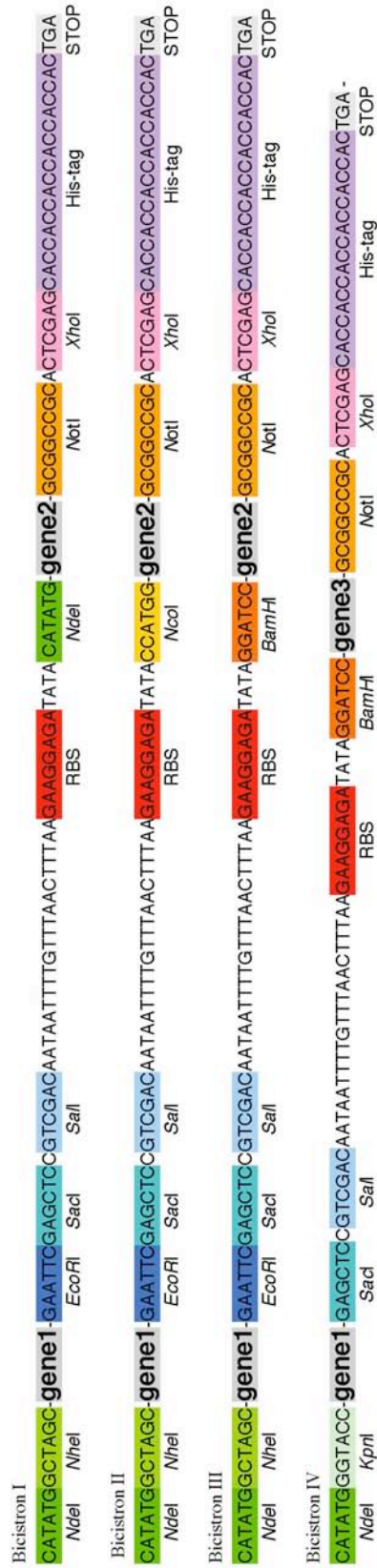
- Trigon, S., Serizawa, H., Conaway, J. W., Conaway, R. C., Jackson, S. P., and Morange, M. (1998). Characterization of the residues phosphorylated *in vitro* by different C-terminal domain kinases. *J Biol Chem* **273**, 6769-6775.
- Tudor, M., Murray, P. J., Onufryk, C., Jaenisch, R., and Young, R. A. (1999). Ubiquitous expression and embryonic requirement for RNA polymerase II coactivator subunit Srb7 in mice. *Genes Dev* **13**, 2365-2368.
- Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., *et al.* (2000). A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* **403**, 623-627.
- Valay, J. G., Simon, M., Dubois, M. F., Bensaude, O., Facca, C., and Faye, G. (1995). The KIN28 gene is required both for RNA polymerase II mediated transcription and phosphorylation of the Rpb1p CTD. *J Mol Biol* **249**, 535-544.
- 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**, 5790-5796.
- Wada, T., Takagi, T., Yamaguchi, Y., Ferdous, A., Imai, T., Hirose, S., Sugimoto, S., Yano, K., Hartzog, G. A., Winston, F., *et al.* (1998a). DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. *Genes Dev* **12**, 343-356.
- Wada, T., Takagi, T., Yamaguchi, Y., Watanabe, D., and Handa, H. (1998b). Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription *in vitro*. *Embo J* **17**, 7395-7403.
- Wallenfang, M. R., and Seydoux, G. (2002). cdk-7 is required for mRNA transcription and cell cycle progression in *Caenorhabditis elegans* embryos. *Proc Natl Acad Sci U S A* **99**, 5527-5532.
- Wang, G., Cantin, G. T., Stevens, J. L., and Berk, A. J. (2001). Characterization of mediator complexes from HeLa cell nuclear extract. *Mol Cell Biol* **21**, 4604-4613.
- Watanabe, Y., Fujimoto, H., Watanabe, T., Maekawa, T., Masutani, C., Hanaoka, F., and Ohkuma, Y. (2000). Modulation of TFIIH-associated kinase activity by complex formation and its relationship with CTD phosphorylation of RNA polymerase II. *Genes Cells* **5**, 407-423.
- Weil, P. A., Luse, D. S., Segall, J., and Roeder, R. G. (1979). Selective and accurate initiation of transcription at the Ad2 major late promoter in a soluble system dependent on purified RNA polymerase II and DNA. *Cell* **18**, 469-484.
- West, M. L., and Corden, J. L. (1995). Construction and analysis of yeast RNA polymerase II CTD deletion and substitution mutations. *Genetics* **140**, 1223-1233.

- 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, 3469-3478.
- Yamaguchi, Y., Takagi, T., Wada, T., Yano, K., Furuya, A., Sugimoto, S., Hasegawa, J., and Handa, H. (1999a). NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell* 97, 41-51.
- Yamaguchi, Y., Wada, T., and Handa, H. (1998). Interplay between positive and negative elongation factors: drawing a new view of DRB. *Genes Cells* 3, 9-15.
- Yamaguchi, Y., Wada, T., Watanabe, D., Takagi, T., Hasegawa, J., and Handa, H. (1999b). Structure and function of the human transcription elongation factor DSIF. *J Biol Chem* 274, 8085-8092.
- Yamamoto, S., Watanabe, Y., van der Spek, P. J., Watanabe, T., Fujimoto, H., Hanaoka, F., and Ohkuma, Y. (2001). Studies of nematode TFIIE function reveal a link between Ser-5 phosphorylation of RNA polymerase II and the transition from transcription initiation to elongation. *Mol Cell Biol* 21, 1-15.
- Yang, Z., Zhu, Q., Luo, K., and Zhou, Q. (2001). The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature* 414, 317-322.
- Yankulov, K. Y., and Bentley, D. L. (1997). Regulation of CDK7 substrate specificity by MAT1 and TFIIH. *Embo J* 16, 1638-1646.
- Yao, S., Neiman, A., and Prelich, G. (2000). BUR1 and BUR2 encode a divergent cyclin-dependent kinase-cyclin complex important for transcription in vivo. *Mol Cell Biol* 20, 7080-7087.
- Yao, S., and Prelich, G. (2002). Activation of the Bur1-Bur2 cyclin-dependent kinase complex by Cak1. *Mol Cell Biol* 22, 6750-6758.
- Yik, J. H., Chen, R., Nishimura, R., Jennings, J. L., Link, A. J., and Zhou, Q. (2003). Inhibition of P-TEFb (CDK9/Cyclin T) kinase and RNA polymerase II transcription by the coordinated actions of HEXIM1 and 7SK snRNA. *Mol Cell* 12, 971-982.
- Young, B. A., Gruber, T. M., and Gross, C. A. (2002). Views of transcription initiation. *Cell* 109, 417-420.
- Yudkovsky, N., Ranish, J. A., and Hahn, S. (2000). A transcription reinitiation intermediate that is stabilized by activator. *Nature* 408, 225-229.
- Zhou, M., Halanski, M. A., Radonovich, M. F., Kashanchi, F., Peng, J., Price, D. H., and Brady, J. N. (2000). Tat modifies the activity of CDK9 to phosphorylate serine 5 of the RNA polymerase II carboxyl-terminal domain during human immunodeficiency virus type 1 transcription. *Mol Cell Biol* 20, 5077-5086.

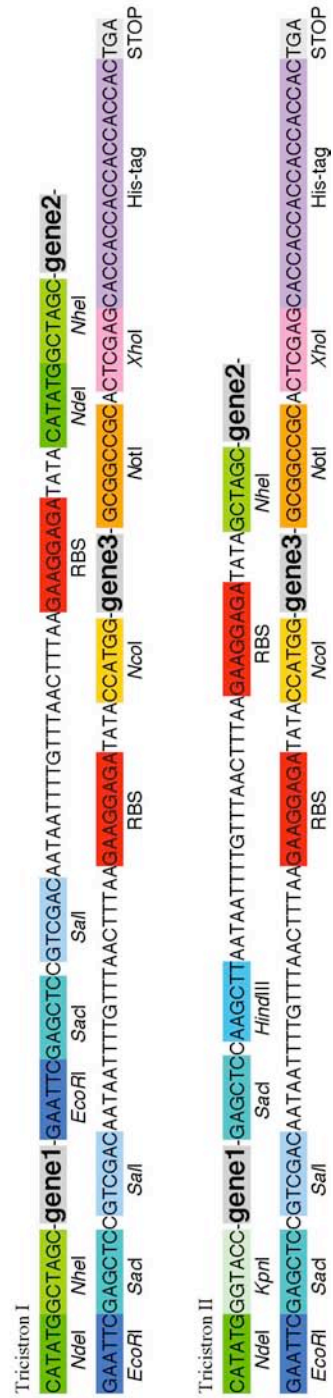
Supplementary Material

Vectors are based on pET21b/24b family – the following can be regarded as insertions between the NdeI and the NotI site of these vectors. Note that the individual gene cloned in the construct might contain additional unwanted cleavage sites.

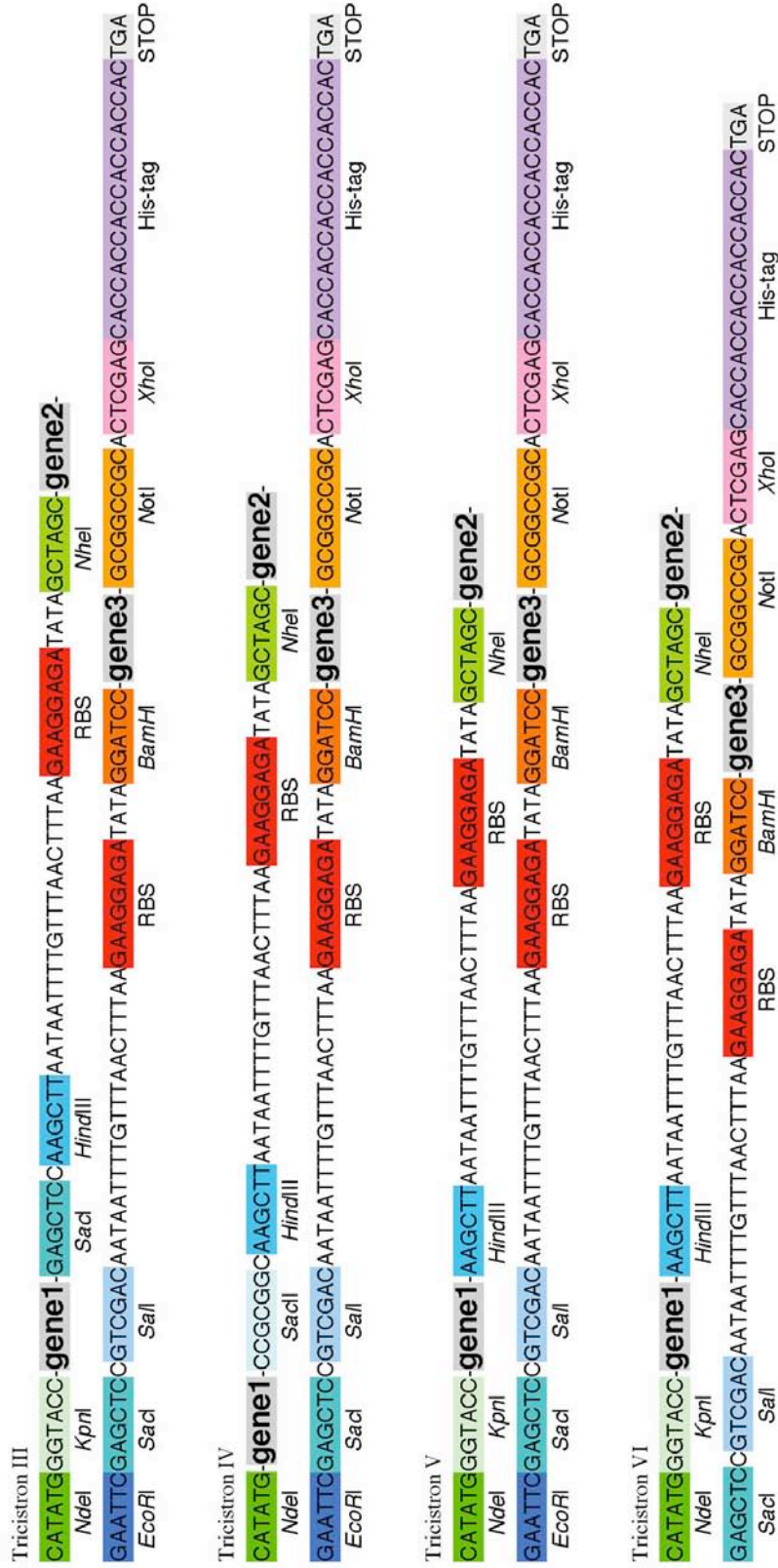
Bicistrons



Tricistrons



Tricistrons – continued



Curriculum Vitae

SABINE HÖPPNER

Personal Data

Date of Birth:	November 29, 1975
Place of Birth:	Mönchengladbach, Germany
Marital Status:	married
Children:	Nicolas Leonid Höppner Simon Christopher Höppner
Citizenship:	Germany

Education

1986-1988	St. Bernhard Gymnasium, Willich
1988-1992	Gymnasium Vaterstetten, Vaterstetten
1992-1993	Institut Lluís Peguera, Manresa, Spain
1993-1995	Gymnasium Vaterstetten, Vaterstetten

Academic Training

1995-1998	LMU Munich, Faculty of Biology, Undergraduate course of Biology
1998-1999	Universidad de Concepción, Chile, Faculty of Biology, Master course of Biology, DAAD scholarship
1999-2000	LMU Munich, Faculty of Biology Undergraduate course of Biology
2000-2001	Diploma student in molecular biology and immunology with Prof. Dr. Rudolf Grosschedl at the Gene Center, LMU Munich
2001-2005	Graduate studies in macromolecular crystallography with Prof. Dr. Patrick Cramer at the Gene Center, LMU Munich