

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

Iwr1 directs RNA polymerase II nuclear import



Elmar Geza Czeko
aus
Temeschburg, Rumänien

2010

Erklärung

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

Ehrenwörtliche Versicherung

Diese Dissertation wurde selbständig und ohne unerlaubte Hilfe erarbeitet.

München, am 14. Dezember 2010

Elmar Czeko

Dissertation eingereicht am 14. Dezember 2010

1. Gutachter: Prof. Dr. Patrick Cramer
2. Gutachter: Prof. Dr. Dietmar Martin

Mündliche Prüfung am 10. Februar 2011

Acknowledgements

Life science research is indeed a social science. Without others, nobody would succeed. Therefore, I was glad to spend my PhD thesis in Patrick Cramer's laboratory at the Gene Center in Munich – undoubtedly a social place.

I want to thank you, *Patrick*, not only for letting me contribute to this environment over the last years, but also for learning from you beyond science. The way you are leading your lab is, as I think, based on trust. Trust in the capabilities, potential and motivation of your people. I am confident to benefit from this experience in my future professional life. Thank you for giving me complete freedom over my project, for the chance to go to diverse seminars and conferences and once to have the honour of representing you as a speaker at a workshop in Barcelona.

Dietmar, I want to thank you for your ongoing interest in my project as well as for your advice. I very much enjoyed our bike excursions; it is just a pity that we did not do them more frequently.

For my project I got a lot of advice and help from people in our lab. Thank you: *Alan* for discussion and first aid crystallographic processing of this tiny and twinned crystal. *Andreas* for having the right feeling about the non-ChIP-ability of Iwr1, we saved a lot of time there and you had the chance to beat the famous Struhl lab. *Anselm* for discussions, help with cryo-EM and computing: You have the bridge, Number 1. *Christian* – whom else – for extinguishing the fire on my bench. Not to forget the talks about politics and the economy. *Claudia* (Blattner) for temporarily sharing the enthusiasm for long-distance runs. *Claudia* (Buchen) for keeping the lab chaos at bay. *Elisabeth* for our informal group meetings and your unmatched expertise. *Gerke* for parental advice. *Kerstin* (Maier) for help with gene expression profiling and yeast. And for finding a way back from Barcelona against all odds. *Laurent* and *Martin* for discussions and for challenging questions. *Rieke* for bestowing some colour upon our bay. *Sarah* for TBP and TFIIB and also for handing over the baton. *Stefan* (Benkert) for fermenting this slowly growing yeast strain. *Stefan* (Jennebach) for Pol I and the occasional quiz question. *Tobias* for always sharing your knowledge, especially when I was new to the lab and for help with R/Bioconductor.

Also I want to thank my students *Christian*, *Julia* and *Simon*. They all helped me to advance my project. Even more, they challenged me to grow into the responsibilities of a supervisor. *Christian*, I want to thank you for your interest and your reliability.

In addition, other people from the Gene Center supported me. Thank you *Achim* for the transcription factor comparison. *Christophe* for introducing me to the fluorescent microscope. *Heidi* and *Lina* for advice on yeast. Also, I have to thank *Atlanta* Cook from the MPI of Biochemistry for the karyopherin proteins and for the private lecture on nuclear traffic.

Meinen Eltern und meiner Oma danke ich für die Unterstützung während meines Studiums und meiner Auslandsaufenthalte. Meinem Vater danke ich dafür, frühzeitig mein Interesse für Naturwissenschaften geweckt zu haben.

Nina, danke für Dein Verständnis und Deinen Rat. Für unsere Familie. Und für *Julius*.

Summary

Pol II is the conserved 12-subunit enzyme that transcribes mRNA from protein-coding genes in the nucleus of eukaryotic cells. Whereas the structure and function of Pol II have been extensively studied, its biogenesis is not well understood and the mechanism of how Pol II is imported into the nucleus remains elusive. This work establishes that the conserved protein Iwr1 specifically binds Pol II between its two largest subunits Rpb1 and Rpb2, and directs its nuclear import.

Iwr1 was previously identified in co-purifications with Pol II subunits and deletion of its gene resulted in a pleiotropic phenotype, suggesting a general cellular function. Moreover, from proteome-wide localization data Iwr1 was known to reside both in the nucleus and in the cytoplasm. While this work was ongoing, Iwr1 was reported to bear a nuclear export sequence (NES), further supporting nucleo-cytoplasmic shuttling of the protein.

Analysis of the Iwr1 sequence revealed an N-terminally located classical bipartite nuclear localization sequence (NLS) that was confirmed to be functional *in vivo* and required for nuclear import of Iwr1. Based on these results the hypothesis was proposed that Iwr1 is involved in nuclear import of Pol II. To validate this, IWR1 was deleted in yeast strains carrying an EGFP-fusion with Rpb1, Rpb3 and Rpb4, respectively. Cytoplasmic accumulation of Pol II subunits was observed, while localization of the Pol I and Pol III largest subunits was unaffected in the absence of Iwr1. Pol II nuclear localization could be re-established upon expression of the full-length protein or of a region of Iwr1 that was sufficient to bind Pol II *in vitro*, while variants with mutations in the bipartite NLS were not functional. The mechanism seems to be conserved from yeast to humans, as a partial rescue of the Pol II delocalization was also possible with the human homologue of Iwr1.

According to a cryo-electron microscopy (EM) map of the Pol II-Iwr1 complex, Iwr1 binds Pol II in the active centre cleft between Rpb1 and Rpb2, thereby possibly sensing complete assembly of the enzyme and thus restricting nuclear import to functional complexes. The cryo-EM map was validated by competition experiments with nucleic acids and general transcription factors, which also led to a proposed mechanism of Iwr1 eviction from the complex as Pol II is recruited to gene promoters in the nucleus.

In agreement with the nucleic acid competition, Iwr1 could not be detected at chromatin in immunoprecipitation (ChIP) experiments. Nuclear extracts from Δ iwr1 cells fully supported *in vitro* transcriptional initiation, which was decreased by addition of recombinant Iwr1. Together with these results a comparison of the Δ iwr1 gene expression profile with the profiles of 263 transcription factor deletions argues against the notion of Iwr1 being a transcription factor.

Finally, an increase of Rpb1 levels and half-life in Δ iwr1 as compared to wild-type cells is described, possibly reflecting an involvement of Iwr1 in Pol II biogenesis as a second function of the protein.

In conclusion, a cyclic model for Iwr1-directed Pol II nuclear import is presented.

Publications

Parts of this work will be published:

Czeko E, Seizl M, Augsberger C, Mielke T, Cramer P (2011)

Iwr1 directs RNA polymerase II nuclear import.

Mol Cell.

Contents

ACKNOWLEDGEMENTS	II
SUMMARY	III
PUBLICATIONS.....	IV
1 INTRODUCTION	1
1.1 EUKARYOTIC RNA POLYMERASES	1
1.2 RNA POLYMERASE II ASSOCIATED FACTORS	2
1.3 IWR1 – INTERACT WITH RNA POLYMERASE II.....	3
1.3.1 <i>Iwr1</i> in <i>Saccharomyces cerevisiae</i>	3
1.3.2 <i>Iwr1</i> homologues in plants.....	5
1.4 THE “LIFE CYCLE” OF RNA POLYMERASES.....	5
1.4.1 <i>Biogenesis</i>	6
1.4.2 <i>Degradation</i>	9
1.5 NUCLEAR TRAFFIC, KARYOPHERINS AND TRANSCRIPTION	12
1.5.1 <i>Nuclear traffic and karyopherins</i>	12
1.5.2 <i>Nuclear import of transcriptional factors</i>	13
1.6 AIMS OF THIS STUDY.....	15
2 MATERIALS AND METHODS	16
2.1 MATERIALS	16
2.1.1 <i>Bacterial strains</i>	16
2.1.2 <i>Yeast strains</i>	16
2.1.3 <i>Oligonucleotides</i>	17
2.1.4 <i>Plasmids</i>	20
2.1.5 <i>Reagents and Consumables</i>	23
2.1.6 <i>Media and additives</i>	23
2.1.7 <i>Buffers and solutions</i>	24
2.2 GENERAL METHODS.....	26
2.2.1 <i>Preparation and transformation of competent bacterial cells</i>	26
2.2.2 <i>Cloning and mutagenesis</i>	27
2.2.3 <i>Protein expression in Escherichia coli</i>	29
2.2.4 <i>Protein analysis</i>	29
2.2.5 <i>Quantitative western blots</i>	31

2.2.6	<i>Bioinformatic tools</i>	32
2.3	YEAST GENETICS AND ASSAYS.....	32
2.3.1	<i>Isolation of genomic DNA from yeast</i>	32
2.3.2	<i>Transformation of yeast cells</i>	32
2.3.3	<i>Gene disruption and epitope-tagging</i>	33
2.3.4	<i>Long-term storage of yeast strains</i>	33
2.3.5	<i>Determination of yeast doubling times</i>	34
2.3.6	<i>Determination of protein half-lives</i>	34
2.3.7	<i>In vitro transcription assays</i>	34
2.4	FLUORESCENCE MICROSCOPY	35
2.4.1	<i>Paraformaldehyde fixation of yeast cells</i>	35
2.4.2	<i>Fluorescence microscopy data collection</i>	36
2.4.3	<i>Analysis of localization ratios</i>	36
2.4.4	<i>Fluorescence microscopy figure preparation</i>	36
2.5	PREPARATION OF THE RNA POLYMERASE II-IWR1 COMPLEX	37
2.5.1	<i>Purification of recombinant Iwr1 and variants</i>	37
2.5.2	<i>Purification of endogenous RNA polymerase II</i>	38
2.5.3	<i>Purification of the recombinant Rpb4/7 subcomplex</i>	39
2.5.4	<i>Assembly of the RNA polymerase II-Iwr1 complex</i>	40
2.5.5	<i>Competition of Iwr1 with nucleic acids and GTFs for RNA polymerase II binding</i>	41
2.5.6	<i>Reciprocal pull-down experiments</i>	41
2.5.7	<i>Static light scattering analysis</i>	42
2.6	STRUCTURAL ANALYSIS OF THE RNA POLYMERASE II-IWR1 COMPLEX	42
2.6.1	<i>Negative stain</i>	42
2.6.2	<i>Preparation of cryo-grids</i>	43
2.6.3	<i>Cryo-electron microscopic data collection</i>	43
2.6.4	<i>Single particle reconstruction</i>	43
2.7	GENE EXPRESSION PROFILING	44
2.7.1	<i>Sample preparation and microarray measurement</i>	44
2.7.2	<i>Statistical data analysis</i>	45
3	RESULTS AND DISCUSSION	46
3.1	RNA POLYMERASE II NUCLEAR LOCALIZATION IS SPECIFICALLY DEPENDENT ON THE IWR1 NLS	46
3.1.1	<i>Iwr1 contains a bioinformatically predicted bipartite NLS that is functional in vivo</i>	46

3.1.2	<i>RNA polymerase II mislocalizes to the cytoplasm in cells lacking Iwr1</i>	48
3.1.3	<i>RNA polymerase II nuclear localization requires the Iwr1 NLS</i>	49
3.1.4	<i>The human homologue of Iwr1 partially rescues RNA polymerase II nuclear localization in Saccharomyces cerevisiae</i>	50
3.1.5	<i>RNA polymerase II nuclear localization is independent of Rpb4</i>	51
3.1.6	<i>RNA polymerase I and III nuclear localization are independent of Iwr1</i>	52
3.2	IWR1 BINDS RNA POLYMERASE II DIRECTLY VIA AN N-TERMINAL BINDING REGION	53
3.2.1	<i>Iwr1 binds directly and specifically to RNA polymerase II</i>	54
3.2.2	<i>Iwr1 contains a minimal RNA polymerase II-binding region that is sufficient for nuclear localization of the enzyme</i>	55
3.2.3	<i>RNA polymerase II-binding is not affected by mutations of the Iwr1 NLS</i>	56
3.3	IWR1 BINDS IN THE RNA POLYMERASE II ACTIVE CENTRE CLEFT	57
3.3.1	<i>A major cryo-EM difference density locates Iwr1 in the RNA polymerase II active centre cleft</i>	57
3.3.2	<i>Competition experiments with nucleic acids confirm the predicted location of Iwr1 on RNA polymerase II</i>	59
3.3.3	<i>Iwr1 is displaced from RNA polymerase II by transcription initiation factors</i>	61
3.3.4	<i>Iwr1 hampers the crystallization of RNA polymerase II</i>	62
3.4	IWR1 DOES NOT SEEM TO BE INVOLVED IN TRANSCRIPTION	63
3.4.1	<i>Iwr1 is not required for promoter-dependent transcription in vitro</i>	63
3.4.2	<i>Iwr1 is not present at chromatin</i>	64
3.4.3	<i>The gene expression profile of the $\Delta iwr1$ strain is dissimilar as compared to profiles of 263 transcription factor gene deletion strains</i>	65
3.5	NUCLEAR IMPORT OF RNA POLYMERASE II IS PRIMARILY DEPENDENT ON IWR1 AND YET REDUNDANT	68
3.5.1	<i>Cellular growth rate is independent of RNA polymerase II nuclear localization above a critical threshold</i>	68
3.5.2	<i>Rpb1 levels and protein half-life are increased in $\Delta iwr1$ cells</i>	70
4	CONCLUSIONS AND OUTLOOK	72
4.1	A MODEL FOR RNA POLYMERASE II NUCLEAR IMPORT	72
4.2	EXTENDING THE PROPOSED MODEL	74
4.3	IWR1 MIGHT INTEGRATE RNA POLYMERASE II BIOGENESIS AND NUCLEAR IMPORT	75
5	REFERENCES	77
	ABBREVIATIONS.....	89
	CURRICULUM VITAE.....	91

1 Introduction

1.1 Eukaryotic RNA polymerases

In bacteria and archaea, a single DNA-dependent RNA polymerase suffices to synthesize all of a cell's transcriptional RNA. Eukaryotes, however, diversified this task and generally possess three different nuclear RNA polymerases, each responsible for synthesis of different classes of RNAs.

RNA polymerase I (Pol I) is located in nucleoli and synthesizes precursors of most ribosomal RNAs. RNA polymerase II (Pol II) acts in the nucleoplasm and synthesizes messenger RNA (mRNA) precursors, small nuclear and nucleolar RNAs and other non-coding RNAs. RNA polymerase III (Pol III) also resides in the nucleoplasm and synthesizes the precursors of 5S ribosomal RNA, transfer RNAs, and other small RNAs. In plants, two additional closely related RNA polymerases, Pol IV and Pol V, exist that are involved in RNA-dependant DNA methylation and heterochromatin formation (Dalmay et al, 2000; Herr et al, 2005).

Eukaryotic RNA polymerases vary in their subunit composition and their molecular weights (Table 1). The RNA polymerase core structure is conserved among all three domains of life (Hirata & Murakami, 2009). Five subunits are shared among all RNA polymerases (Rpb5, 6, 8, 10 and 12) and two between Pol I and Pol III (AC40 and 19), whereas all others are specific. Not all subunits are essential for transcriptional elongation, such as Rpb4, 7 and 9 in the case of Pol II.

Although a plethora of mechanistic and structural details on eukaryotic RNA polymerases is described (Cramer et al, 2008), the understanding of their "life cycle" beyond transcription, namely their biogenesis, nuclear translocation and degradation is only about to be developed (Cloutier & Coulombe, 2010).

Table 1 · Subunits of RNA polymerases

RNA polymerase	Pol I	Pol II	Pol III	bacterial
core subunits – largest/specific	A190	Rpb1	C160	β'
	A135	Rpb2	C128	β
– partially shared/ α -motif containing	AC40	Rpb3	AC40	α
	AC19	Rpb11	AC19	α
– specific	A12.2*	Rpb9*	C11	
– shared	ABC27	ABC27 (Rpb5)	ABC27	
	ABC23	ABC23 (Rpb6)	ABC23	ω^*
	ABC14.5	ABC14.5 (Rpb8)	ABC14.5	
	ABC10 β	ABC10 β (Rpb10)	ABC10 β	
	ABC10 α	ABC10 α (Rpb12)	ABC10 α	
Rpb4/7 subcomplex	A14*	Rpb4*	C17	
	A43	Rpb7	C25	
TFIIF-like subcomplex ¹	A49*	Tfg1	C37	
	A34.5*	Tfg2	C53	
Pol III-specific subcomplex			C82	
			C34	
			C31	
Number of subunits	14	12	17	5
Molecular weight [kDa]	589	514	693	~400

¹(Geiger et al, 2010), *subunit is not essential in YPD/LB medium in *Saccharomyces cerevisiae*/*Escherichia coli*

1.2 RNA polymerase II associated factors

The central enzyme in eukaryotic transcription of protein-coding genes is Pol II. To act in a regulated manner, Pol II is dependent on a series of factors. Most prominent are the general transcription factors (GTFs, TFIIA, B, D, E, F, H) that form the pre-initiation complex together with Pol II and that are sufficient for basal transcription (Hahn, 2004; Thomas & Chiang, 2006). For activator-dependent transcription additional cofactors such as the Mediator complex or TBP-associated factors (TAFs) are required. Furthermore, while transcription proceeds Pol II is bound by factors involved in elongation, co-transcriptional RNA processing, DNA repair and termination.

Pol II in association with TFIIF is one of the major stable Pol II complexes in yeast (Rani et al, 2004). About half of the polymerase is present in this form. Yeast TFIIF is made up of three subunits (Tfg1, 2 and 3) and is involved at multiple steps in initiation such as recruiting Pol II to the promoter (Flores et al,

1991), selection of the transcription start site together with TFIIB (Sun & Hampsey, 1995), entry of TFIIE and TFIIH into the PIC (Flores et al, 1992), DNA strand separation together with TFIIE (Coulombe & Burton, 1999) and promoter escape (Yan et al, 1999). Also several functions of TFIIF in elongation are described (Thomas & Chiang, 2006).

Only recently, mechanistic insight into transcriptional initiation has been gained with an atomic structure of TFIIB in complex with the polymerase (Kostrewa et al, 2009). The structure rationalizes DNA opening, scanning for the transcription start site with a TFIIB region called the B-reader, transition to RNA elongation, release of TFIIB from Pol II and finally elongation complex formation. Structurally, TFIIB extends from the Pol II dock domain via the RNA exit tunnel to the hybrid-binding site and active centre, where the B-reader is located, to the rudder and clamp coiled-coil, and to the Pol II wall.

Thus, Pol II binding factors are key to understand transcription. To better understand polymerase biology away from DNA, factors involved in these processes need to be identified and investigated.

1.3 Iwr1 – Interact with RNA polymerase II

1.3.1 Iwr1 in *Saccharomyces cerevisiae*

In an attempt to comprehensively describe protein complexes in *Saccharomyces cerevisiae* by a genome-wide tandem-affinity purification (TAP) and mass spectrometry (MS) identification approach an uncharacterized protein encoded by the ORF YDL115c was found to be co-purified with TAP-tagged Rpb3 (Gavin et al, 2002). Most other Pol II subunits as well as the TFIIF subunits Tfg1 and Tfg2 were identified in the same co-purification. Accordingly, the protein was later named *Interact With RNA polymerase II 1* (Iwr1). The association of Iwr1 with Pol II and TFIIF was confirmed by a second genome-wide TAP/MS-study (Krogan et al, 2006).

In an earlier screen the deletion of YDL115c was found to cause a pleiotropic phenotype, suggesting a general function of the gene (Bianchi et al,

2001). Moreover, the deletion strain showed an increased sensitivity to the antifungal protein K1 killer toxin (Page et al, 2003). In a study using synthetic genetic arrays (SGA) that concentrated on around 800 genes involved in chromatin biology, profiles were assigned to a certain gene that were made up of measured colony sizes of all double knock-outs of that gene with each of the other genes under investigation. Based thereon, the IWR1 profile was found to be most similar to the profiles of the Mediator components MED20 and MED31, suggesting a critical role for properly regulated Pol II transcription (Collins et al, 2007).

Due to an unannotated intron in the IWR1 gene the start of the protein-coding region in the databases was initially incorrect. Although the intron was predicted bioinformatically (Brachat et al, 2003) the annotation was corrected only after experimental prove became available (Juneau et al, 2007).

Recently, IWR1 was found in a screen for suppressors of negative co-factor 2 (NC2) requirement (Peiro-Chova & Estruch, 2009). In the same study, the fusion protein LexA-Iwr1 was shown to be competent in transcriptional activation in a strain carrying the HIS3 and lacZ reporters that both included a LexA binding site in their promoter. Also, ChIP/RT-PCR experiments with a TAP-, HA- and myc-tagged version of Iwr1 were performed on promoters of genes (ARG1, PHO5, SUC2) that showed the highest fold-change in a gene expression profile of the $\Delta iwr1$ strain (compared to wild type). Although the presence of Iwr1 at the promoter remained unconfirmed, Iwr1 was suggested to be a transcription factor.

Iwr1 possesses a leucine-rich nuclear export sequence (NES) (Peiro-Chova & Estruch, 2009). Upon deletion of the NES or in a temperature sensitive Xpo1 nuclear export factor strain under non-permissive conditions, Iwr1-GFP shows an altered localization that changes from pan-cellular in wild type cells to nuclear in the mutants.

1.3.2 Iwr1 homologues in plants

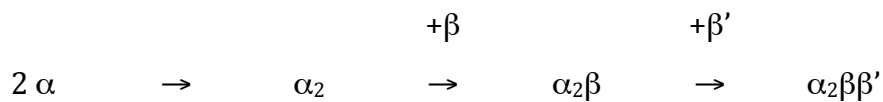
In *Arabidopsis thaliana* DMS4, the homologue to yeast IWR1, was found in a screen for factors involved in RNA-directed DNA methylation (RdDM) in a developmental context (Kanno et al, 2010). Unlike other mutants defective in RdDM, *dms4* mutants have a pleiotropic developmental phenotype. DMS4 cDNA partially complements the K1 killer toxin hypersensitivity of the yeast Δ *iwr1* strain, suggesting some functional conservation. Mutations in DMS4 directly or indirectly affect Pol IV-dependent secondary siRNAs, Pol V-mediated RdDM, Pol V-dependent synthesis of intergenic noncoding RNA and altered expression of many Pol II driven genes. A second study in *Arabidopsis thaliana* made similar observations. There, RDM4, identical with DMS4, is additionally shown to physically interact with Pol V (He et al, 2009). Therefore, DMS4 may be a regulatory factor for multiple RNA polymerases in plants. However, Pol IV/V are more similar to Pol II, than to Pol I or Pol III (Herr et al, 2005; Kanno et al, 2005).

1.4 The “life cycle” of RNA polymerases

Although a detailed understanding about the structural and functional aspects of RNA polymerases in general and Pol II in particular has been reached (Cramer et al, 2008), much remains to be learnt about the “life cycle” of these intricate cellular machines. Their biogenesis, nuclear translocation and degradation are poorly or at best fragmentarily investigated. Here, a comprehensive review is attempted, summing up the current knowledge on RNA polymerase biogenesis and degradation.

1.4.1 Biogenesis

Already in the early seventies the first studies on *Escherichia coli* RNA polymerase assembly were performed (Ishihama & Ito, 1972). After complete dissociation of bacterial RNA polymerase in 6 M urea, subunits were separated by successive column chromatography and binary combinations for reassembly were tested. A significant interaction was observed only between the α and β subunits, and the complex formed was found to have the composition $\alpha_2\beta$. Addition of the dialysed β' subunit converted this complex into the complete enzyme, suggesting that the assembly of bacterial RNA polymerase is a sequential process:



Later on, the *in vitro* assembly pathway was confirmed *in vivo* with the use of pulse-labelling experiments, temperature sensitive core subunit mutants and the identification of assembly intermediates in cell extracts (Ishihama, 1981).

Due to the high structural conservation between eukaryotic and bacterial RNA polymerases (Cramer et al, 2001), these findings later turned out to be a valid starting point for eukaryotic RNA polymerase assembly. Dissociation experiments with yeast Pol II yielded an Rpb2/3/11 subassembly that was stable at 6 M urea (Kimura et al, 1997). At a lower concentration of 4 M urea a substoichiometric level of Rpb10 was bound to the Rpb2/3/11 subassembly, while substoichiometric levels of Rpb8 were bound to Rpb1. The use of temperature sensitive assembly mutants of Rpb1, Rpb2 and Rpb3 led to the identification of Rpb2/3 as an early subcomplex in pulse-labelling experiments (Kolodziej & Young, 1991). A western blot based quantification of all Pol II subunits in *Schizosaccharomyces pombe* revealed that Rpb3 is least abundant and limiting for complex formation (Kimura et al, 2001). The two largest subunits and Rpb7 were present at up to two-fold higher levels. Rpb4 was eleven times,

Rpb9 and Rpb11 were four times, and the common subunits (Rpb5, 6, 8, 10 and 12) were seven to fifteen times more abundant than Rpb3. Again, the tight association of Rpb3 and Rpb2 could be confirmed, as overexpressed Rpb3 ended up in complex with Rpb2. Smaller subunits (Rpb4 to Rpb12) were also identified in unassembled form by glycerol gradient centrifugation, meaning that they can exist at least temporarily in free form in the cell. With Rpb1, Rpb2 and Rpb3/11 being the homologues of the bacterial β' , β and α subunits, respectively, the yeast subcomplexes are in agreement with the assembly pathway derived for bacteria.

The common subunits of Pol I and Pol III, AC40 and AC19, which share the bacterial α motif with Rpb3 and Rpb11 were found upon mutation to be extragenetically suppressed by the shared subunit ABC10 β (Rpb10) (Lalo et al, 1993). This is in agreement with Rpb10 binding to the Rpb2/3/11 subassembly and could be an indication of a conserved assembly pathway among eukaryotic RNA polymerases.

Another shared subunit, ABC10 α (Rpb12), is unique in the respect, that diploid cells containing only one functional gene copy have a partial growth defect (Rubbi et al, 1999). Preventing *de novo* expression of ABC10 α leads to a parallel decrease in all RNA polymerase transcription followed by complete growth arrest. A certain ABC10 α double mutant exhibits a temperature sensitive phenotype at 37 °C which is characterized by declining Pol III levels and impaired Pol III transcription of tRNAs. As the specific activity and heat stability of Pol III with wild-type and mutated ABC10 α are identical and considering the ABC10 α requirement for Pol II, and I a critical role of ABC10 α in the assembly of all three RNA polymerases has been proposed.

The common subunit Rpb6 and its bacterial homologue ω have been implicated in RNA polymerase assembly and stability (Minakhin et al, 2001). A “latching mechanism” has been proposed in which Rpb6/ ω stabilizes the N- and C-terminus of the largest subunit, facilitating its association with the Rpb2/3/11 or $\alpha_2\beta$ assembly intermediate, respectively. Overexpression of Rpb6 suppresses the phenotype of the temperature sensitive Rpb1-1 mutant (Nonet et al, 1987) that is characterized by immediate transcriptional shutdown at elevated temperatures, underpinning the relevance of Rpb6 for the stability of the formed

complex. A comparison of protein levels between different growth phases uncovered that absolute Rpb6 levels decreased to only about 60% in stationary phase as compared to exponential phase, while all other Pol II subunits decreased to 40 to 50%, which was interpreted in light of its alleged role in assembly (Sakurai & Ishihama, 2002).

In earlier studies dedicated to the *assembly* of Pol II, the distinction between the structure, i.e. the final spatial configuration of the complex, and its biogenesis, i.e. the process of building the complex over time, was blurred. Pair-wise interactions between Rpb subunits were probed with far-western and cross-linking experiments, that led to structural insight but left question of biogenesis unanswered (Ishiguro et al, 1998). By these means, the Rpb3/11 *heterodimer* of the core polymerase was identified which was distinct from the homologous α_2 *homodimer* in bacterial RNA polymerase (Yasui et al, 1998). A core Pol II backbone model finally answered all remaining structural questions and provided a framework to validate subunit interaction data available at that time (Cramer et al, 2000). Importantly, the possibility of an Rpb3/11/10 assembly intermediate that was tightly associated with the second largest subunit Rpb2 was structurally confirmed. Also, the critical role of the shared subunit ABC10 α (Rpb12) (Rubbi et al, 1999) was rationalized by the core Pol II backbone model, as Rpb12 bridges between Rpb2 and the Rpb3/11/10 subassembly. By stabilizing the interaction between the second largest and α -motif-containing subunits, ABC10 α appears to play an essential role in the assembly or maintenance of all eukaryotic RNA polymerases.

Biogenesis of eukaryotic RNA polymerases probably involves more players than the subunits making up the final enzyme. Attempts to recombinantly express Pol II failed (Kimura & Ishihama, 2000). Expression trials in insect cells did not yield functional Pol II, little, if any, complete complexes could be obtained and mostly subcomplexes without the two largest subunits were purified. The reason seems to lie in the requirement for chaperone assistance during the folding of the largest subunits. As published during the course of this work and inferred from a mass spectrometry-based quantitative proteomic approach HSP90 activity stabilizes incompletely assembled Rpb1 in the cytoplasm of human cells (Boulon et al, 2010). Further, the HSP90 co-

chaperone hSpagh (RPAP3) as well as the R2TP/prefoldin-like complex bind to a cytoplasmic complex of Rpb1 and Rpb8, confirming this assembly intermediate (Kimura et al, 1997). hSpagh (RPAP3) was also found to bind free Rpa194 of Pol I, suggesting a general role in assembling RNA polymerases. In agreement with previous data (Cramer et al, 2000; Kimura et al, 1997), a cytoplasmic Rpb2/3/11/10/12 subcomplex was identified. Importantly, the site of Pol II assembly was described to be the cytoplasm. Depletion of any Pol II subunit resulted in cytoplasmic accumulation of Rpb1, which was resistant to proteolytic degradation after addition of α -amanitin. As α -amanitin specifically promotes degradation of the elongating form of Pol II (Nguyen et al, 1996), Rpb1 needs to correspond to newly synthesized protein and therefore Pol II assembly entirely has to take place in the cytoplasm and requires the presence of all polymerase subunits.

In summary, there seems to be a consistent understanding on the step-wise association of polymerase subunits and the assembly intermediates formed. However, to which extent these associations are regulated by additional factors and at which steps the assistance of chaperones is needed remains elusive. Ultimately, a structural understanding of defined assembly intermediates in complex with auxiliary factors is needed.

1.4.2 Degradation

Pol II subunits have individual half-lives, with estimates ranging from 22 min for Rpb1 to 87 min for Rpb9, while the shared subunits Rpb6, Rpb8 and Rpb12 were described as stable (within the course of the measurement) (Belle et al, 2006). In terms of protein degradation, these individual half-lives suggest an individual fate of each subunit as opposed to proteolysis of the Pol II complex as a whole. How this *basal* turnover of Pol II subunits is achieved and regulated has not been investigated in detail. Most studies focused on targeted Pol II degradation after DNA damage or treatment of cells with transcriptional inhibitors such as α -amanitin, actinomycin D or 6-azauracil (Beaudenon et al,

1999; Harreman et al, 2009; Nguyen et al, 1996; Ribar et al, 2007; Somesh et al, 2005; Somesh et al, 2007; Woudstra et al, 2002).

Treatment of cells with α -amanitin led to specific degradation of Rpb1, while the shared subunits Rpb5 and Rpb8 were almost unaffected (Nguyen et al, 1996). The dependence of degradation on elongation was inconclusive as the transcriptional inhibitor actinomycin D abolished α -amanitin induced Rpb1 degradation as opposed to another inhibitor, 5,6-dichloro-1- β -D-ribofuranosyl-benzimidazole (DRB), that did not.

When Pol II encounters a DNA lesion, the damage is either immediately repaired by transcription-coupled nucleotide excision repair (TCR) or Pol II (Rpb1) is degraded to allow for other repair pathways, such as general (global) genome repair to take place (Selth et al, 2010). In mammalian cells, defective TCR gives rise to the severe human disorder Cockayne's syndrome (CS). A link between TCR and DNA damage-induced Pol II degradation is provided by the factor Def1 that forms a complex with the TCR factor Rad26/CSB (Woudstra et al, 2002). Rad26 promotes TCR in conjunction with CSA and delays Def1 induced Pol II degradation. Deletion of Def1 in yeast leads to impaired DNA damage-induced Pol II degradation and to increased 6-azauracil sensitivity that is comparable to DST1 (TFIIS) deletion cells. In an *in vitro* reconstituted system it was shown that Def1 binds to Pol II in a DNA damage-dependent manner and that it is directly involved in ubiquitination of Pol II (Reid & Svejstrup, 2004).

Two ubiquitin ligases have been found to be responsible for ubiquitination of Pol II. The first, Rsp5, is an essential HECT domain E3 ubiquitin ligase that was shown to bind the Rpb1 CTD from ³⁵S-labeled yeast extracts (Huibregtse et al, 1997) and to ubiquitinate Rpb1 upon UV-induced DNA damage (Beaudenon et al, 1999). Basal levels of Rpb1 ubiquitination were found in normally growing cells, suggesting that this specific degradation pathway might also play a role in constitutive turnover of the largest subunit. *In vitro* arrested Pol II elongation complexes were shown to be the preferred substrates for Rpb1 ubiquitination by Rsp5 with further stimulation by Def1 (Somesh et al, 2005). Serine 5 phosphorylation of the Pol II CTD, that is typical for initiating Pol II, inhibited ubiquitination, while Serine 2 phosphorylation, providing a mark for elongating Pol II, did not. Even deceleration of elongation by 6-azauracil

augmented Rpb1 labelling with ubiquitin. This led to the general view, that arrested or paused Pol II is preferentially targeted for ubiquitination and that immobile polymerases at the promoter are protected by a serine 5 phosphorylation mark on the Rpb1 CTD. Mechanistically, each of the two Pol II ubiquitination sites on Rpb1 (K330 and K695) is bound by one Ubc5 protein (a ubiquitin conjugating enzyme, E2) until ubiquitin transfer is triggered when both Ubc5 proteins simultaneously form a complex with an Rsp5 dimer bound to the Rpb1 CTD (Somesh et al, 2007).

The second E3 ubiquitin ligase acting on Pol II is the elongin-cullin complex (Elc1, Ela1, Cul3, Roc1 in yeast), which has been shown to promote Rpb1 lysine 48-linked polyubiquitination (Ribar et al, 2007) – the mark for 26S proteasomal degradation. The distinct roles of the two Pol II ubiquitin ligases, Rsp5 and the elongin-cullin complex have been unravelled lately (Harreman et al, 2009). Rsp5 either leads to mono- or lysine 63-linked polyubiquitination that is trimmed to mono-ubiquitination by the ubiquitin protease Ubp2. Based on this mono-ubiquitin moiety the elongin-cullin complex catalyzes formation of an ubiquitin chain linked via lysine 48, which triggers proteolysis. This cooperation of two sequentially acting ubiquitin ligases possibly forms a checkpoint to ensure Rpb1 degradation only when it is indeed required.

Although intricate mechanistic insights have been obtained about scheduled proteolysis of Rpb1, open questions remain. Does targeting for degradation also take place away from DNA? Are non-degraded subunits recycled and re-assembled into new complexes in the cytoplasm, as has recently been suggested (Boulon et al, 2010)? To put it more broadly, in which ways and via which nuclear import and export pathways are Pol II biogenesis and degradation intertwined?

1.5 Nuclear traffic, karyopherins and transcription

1.5.1 Nuclear traffic and karyopherins

All traffic into and out of the nucleus passes through nuclear pore complexes (NPCs) (Chook & Blobel, 2001; Mosammaparast & Pemberton, 2004; Pemberton & Paschal, 2005). These large multiprotein assemblies with an approximate molecular weight of 60 MDa in yeast and 125 MDa in mammalian cells restrict the size of macromolecules that can freely diffuse between the nucleus and the cytoplasm. Above a molecular weight of around 40 kDa proteins need to be recognized as cargo by karyopherins that mediate transport through the NPC by interacting with its nucleoporins. While yeast cells contain 14 karyopherins, in human cells this family comprises at least 20 members. Most karyopherins are either importins or exportins, the name reflecting the direction in which cargo is translocated relative to the nucleus.

The best characterized adaptor protein is importin- α /importin- β (Kap60/Kap95 in yeast), which is special in the respect that importin- β does not bind its cargo directly, like other karyopherin- β family members, but rather needs importin- α as an adaptor protein. Importin- α recognizes and binds to the so-called basic or classical nuclear localization sequence (NLS) that has initially been characterized from the SV40 large T antigen and from nucleoplasmin. This NLS is either mono- or bipartite, meaning that it consists of either one or two clusters of basic amino acids separated by a linker. The classical NLS has recently been characterized by systematic mutations and measurement of the resultant signal strength in import assays (Kosugi et al, 2009a) providing quantitative data for a bioinformatic NLS identification program that goes beyond simple motif search (Kosugi et al, 2009b). The NLS recognized by other karyopherin- β s are harder to define, like the relatively large M9 NLS, which consists of 38 a. a., is glycine-rich and deficient in basic amino acids (Pollard et al, 1996). In other cases, the three-dimensional structure of the protein seemed to be critical (Rosenblum et al, 1998).

The best characterized exportin is Crm1 (Xpo1 in yeast), which recognizes leucine rich nuclear export sequences (NES) that are utilized by all eukaryotes (Fornerod et al, 1997; Stade et al, 1997). These hydrophobic NES sequences are loosely conserved containing three to four hydrophobic residues in close proximity. Like the classical NLS they can be identified by bioinformatic means, though are harder to predict (la Cour et al, 2004).

Directionality of nuclear transport depends on a RanGTP gradient that is characterized by a high concentration in the nucleus and a low concentration in the cytoplasm (Chook & Blobel, 2001; Mosammaparast & Pemberton, 2004; Pemberton & Paschal, 2005). Importins release their cargo upon RanGTP binding in the nucleus, leading to recycling of the importin to the cytoplasm. There, the intrinsic GTPase activity of Ran is stimulated by Ran GTPase activating protein (RanGAP), leading to hydrolysis of GTP to GDP and release of RanGDP from the importin. RanGDP is then transported back to the nucleus by its proper import factor NTF2. GDP is exchanged for GTP by Ran guanine nucleotide exchange factor (RanGEF), which is bound to chromatin, maintaining a high RanGTP concentration in the nucleus. Contrarily to importins, exportins need RanGTP to bind their cargo in a ternary complex. In the cytoplasm they release their cargo upon Ran GTP hydrolysis.

1.5.2 Nuclear import of transcriptional factors

Research on nuclear transport already identified routes of several transcriptional factors into the nucleus. TFIIA has been shown to be imported by Kap122 (Titov & Blobel, 1999), TFIIS by Kap119 (Albertini et al, 1998), TFIIB and TBP by Kap114 (Hodges et al, 2005; Morehouse et al, 1999) and TFIIF by Kap104 (Suel & Chook, 2009). None of these karyopherins is essential in yeast, although all of their cargoes listed here, with the exception of TFIIS, are, implicating a network of redundancy in import. Partly, this redundancy comes from promiscuous binding of karyopherins to different NLS sequences. In addition, a second level of redundancy is imaginable at the level of cargo, where

proteins undergo transient interactions with each other and thereby are co-imported into the nucleus.

Interestingly, no NLS has been identified in any of the Pol II core subunits. Neither was a localisation signal or a karyopherin described for the Rpb4/7 heterodimer that was reported to undergo nucleo-cytoplasmic shuttling linked to a role in mRNA decay in the cytoplasm (Goler-Baron et al, 2008; Lotan et al, 2005).

Recently, a model was proposed in which the two largest subunits of Pol II are supposed to be imported separately from the smaller ones (Forget et al, 2010). Based on a systematic protein affinity purification and mass spectrometry identification approach a protein interaction network for Pol II in the soluble fraction of human extracts was generated. One of the Pol II interacting factors, the GTPase RPAP4/GPN1 was shown to shuttle between the nucleus and the cytoplasm. Depletion of the factor by siRNA resulted in cytoplasmic accumulation of Rpb1 and Rpb2, while the smaller subunits were unaffected. Treatment of the cells with the CRM1 exportin specific inhibitor leptomycin B trapped RPAP4/GPN1 in the nucleus and also led to cytoplasmic accumulation of the two largest subunits. The essential RPAP4/GPN1 homologue in yeast, Npa3, also bound Pol II and mutation of its 3 a. a. loop resulted in either lethality or slow growth and cytoplasmic mislocalization of Rpb1 and Rpb2. Further, NP3 mutants were hypersensitive to benomyl, an inhibitor of microtubule assembly. This finding might indicate the involvement of microtubules in Pol II nuclear import as it has been shown for other nuclear proteins (Pouton et al, 2007).

However, RPAP4/GPN1 was reported to bind both Rpb1 and the Rpb2/3/10/11/12 assembly intermediate (Boulon et al, 2010), making an earlier involvement in Pol II biogenesis and before nuclear import likely. Moreover, a separate import mechanism for the two largest subunits is not in good agreement with the well established Pol II assembly intermediates (Boulon et al, 2010; Kimura et al, 1997; Kimura et al, 2001; Kolodziej & Young, 1991; Minakhin et al, 2001). Import of subcomplexes also contradicts the requirement for every single Pol II subunit to be present in the cytoplasm for Rpb1 import in the context of the Pol II complex to occur (Boulon et al, 2010).

1.6 Aims of this study

The identification of novel Pol II-interacting factors always bears the promise to obtain new insights into RNA polymerase biology. In this respect, the *Saccharomyces cerevisiae* protein encoded by the ORF YDL115c deserved special attention for two reasons. Firstly, as compared to most of the general transcription factors that on their own only weakly interact with Pol II, it forms a stable complex with the enzyme (Gavin et al, 2002; Krogan et al, 2006). Nevertheless *Interact With RNA polymerase II 1* (Iwr1), as the protein encoded by YDL115c is named today, was described rather late in the history of RNA polymerase research. Secondly, deletion of the IWR1 gene resulted in pleiotropic effects, implying a basic cellular function (Bianchi et al, 2001; Page et al, 2003).

In this study, a both structural and functional approach was chosen to investigate Iwr1. Starting with a bioinformatic analysis of the Iwr1 sequence, a potential N-terminal nuclear localisation sequence (NLS) was identified (see 3.1.1), which together with a published Iwr1 nuclear export sequence (NES) (Peiro-Chova & Estruch, 2009) and Iwr1 localization data from a proteome-wide approach (Huh et al, 2003) suggested nucleo-cytoplasmic shuttling of the protein. Given this shuttling and tight binding of Iwr1 to Pol II (see 3.2.1), a hypothesis was drafted, in which Iwr1 was proposed to be the central and specific factor for nuclear import of Pol II. Therefore, the dependency of the *in vivo* localization of all three eukaryotic RNA polymerases on Iwr1 and variants had to be assessed, both qualitatively and quantitatively in correlation with cellular growth rates (see 3.1). Functional *in vivo* data was integrated with the *in vitro* delineation of an Iwr1 region sufficient for Pol II binding (see 3.2.2). Structurally, the architecture of the Pol II-Iwr1 complex had to be elucidated by cryo-EM (see 3.3.1) and was supposed to be confirmed by competition experiments with nucleic acids (see 3.3.2) and general transcription factors (see 3.3.3). The notion of Iwr1 being a transcription factor (Peiro-Chova & Estruch, 2009) had to be challenged and repudiated by *in vitro* transcription initiation assays (see 3.4.1) and by gene expression analysis and extensive profile comparison (see 3.4.3).

2 Materials and Methods

2.1 Materials

2.1.1 Bacterial strains

Bacterial strains used in this study are listed in Table 2. XL-1 Blue cells were used for cloning purposes and BL21-CodonPlus (DE3) RIL cells for recombinant protein expression.

Table 2 · *Escherichia coli* strains

Strain	Genotype	Source
XL-1 Blue	rec1A, endA1, gyrA96, thi-1, hsdR17, supE44, relA1, lac[F' proAB lacIqZΔM15 Tn10(Tetr)]	Stratagene
BL21-CodonPlus (DE3) RIL	B, F ⁻ , ompT, hsdS(r ⁻ m ⁻), dcm ⁺ , Tetr, gal λ(DE3), endA, Hte [argU, ileY, leuW, Camr]	Stratagene

2.1.2 Yeast strains

Yeast strains used in this study are listed in Table 3. They were generated as part of this work, obtained from Euroscarf or received from other laboratories. Some stains were used for experiments that can be found in Elmar Czeko's laboratory documentation and that are not discussed in this thesis (PCY IDs 233, 234, 247 and 248).

Table 3 · *Saccharomyces cerevisiae* strains. PCY IDs (Patrick Cramer Yeast Identifiers) refer to the laboratory collection.

PCY ID	Strain	Genotype	Source
047	wild-type (haploid)	BY4742; MATa his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0	Euroscarf Y10000
068	wild-type (diploid)	BY4743; MATa/α his3Δ0/his3Δ0 leu2Δ0/leu2Δ0 met15Δ0/MET15 lys2Δ0/LYS2 ura3Δ0/ura3Δ0	Euroscarf Y20000
231	<i>iwr1Δ</i>	BY4742; iwr1::kanMX4	this work
232	<i>iwr1Δ, rpb4Δ</i>	BY4742; iwr1::kanMX4 rpb4::natMX4	this work

PCY ID	Strain	Genotype	Source
233	<i>tfg2Δ</i>	BY4742; <i>MATa</i> or <i>α lys2Δ0</i> or <i>LYS2 met15Δ0</i> or <i>MET15 tfg2::natMX4</i> ; pRS316: <i>URA3 TFG2</i>	this work
234	<i>iwr1Δ, tfg2Δ</i>	BY4742; <i>MATa</i> or <i>α lys2Δ0</i> or <i>LYS2 met15Δ0</i> or <i>MET15 iwr1::kanMX4 tfg2::natMX4</i> ; pRS316: <i>URA3 TFG2</i>	this work
235	<i>Rpb1-EGFP</i>	BY4742; <i>rpb1::RPB1-EGFP-HIS3MX6</i>	this work
236	<i>Rpb1-EGFP, iwr1Δ</i>	BY4742; <i>rpb1::RPB1-EGFP-HIS3MX6 iwr1::kanMX4</i>	this work
237	<i>Rpb3-EGFP</i>	BY4742; <i>rpb3::RPB3-EGFP-HIS3MX6</i>	this work
238	<i>Rpb3-EGFP, iwr1Δ</i>	BY4742; <i>iwr1::kanMX4 rpb3::RPB3-EGFP-HIS3MX6</i>	this work
239	<i>Rpb3-EGFP, rpb4Δ</i>	BY4742; <i>rpb3::RPB3-EGFP-HIS3MX6 rpb4::natMX4</i>	this work
240	<i>Rpb3-EGFP, iwr1Δ, rpb4Δ</i>	BY4742; <i>rpb3::RPB3-EGFP-HIS3MX6 iwr1::kanMX4 rpb4::natMX4</i>	this work
241	<i>Rpb4-EGFP</i>	BY4742; <i>rpb4::RPB4-EGFP-HIS3MX6</i>	this work
242	<i>Rpb4-EGFP, iwr1Δ</i>	BY4742; <i>rpb4::RPB4-EGFP-HIS3MX6 iwr1::kanMX4</i>	this work
243	<i>Rpa190-EGFP</i>	BY4742; <i>rpa190::RPA190-EGFP-HIS3MX6</i>	this work
244	<i>Rpa190-EGFP, iwr1Δ</i>	BY4742; <i>rpa190::RPA190-EGFP-HIS3MX6 iwr1::kanMX4</i>	this work
245	<i>C160-EGFP</i>	BY4742; <i>rpo31::RPO31-EGFP-HIS3MX6</i>	this work
246	<i>C160-EGFP, iwr1Δ</i>	BY4742; <i>rpo31::RPO31-EGFP-HIS3MX6 iwr1::kanMX4</i>	this work
066	Pol II purification strain	<i>MATa</i> or <i>α</i> ; <i>ura3-52 trp1Δ leu2Δ1 his3Δ200 pep4::HIS3 prb1Δ1.6R can1Δ GAL rpb3::URA3-N-6xHis-RPB3</i>	Kashlev lab
247	Pol II purification strain, <i>iwr1Δ</i>	PCY066; <i>iwr1::kanMX4</i>	this work
075	Pol III purification strain	<i>MATα ret1::HIS3 ura3Δ trp1Δ leu2Δ his3Δ met4Δ lys20Δ ade20Δ can10Δ cyh2R</i> ; pNZ85: <i>TRP1 N-(His)6-(FLAG)4-RET1</i>	Bartholomew lab (Zecherle et al, 1996)
248	Pol III purification strain, <i>iwr1Δ</i>	PCY075; <i>iwr1::kanMX4</i>	this work

2.1.3 Oligonucleotides

Oligonucleotides (Table 4) used in this study were ordered at Thermo Scientific in RP-HPLC quality.

Table 4 · Oligonucleotides. The name generally indicates the targeted gene and its species of origin, direction of the oligonucleotide (forward F, reverse R), restriction sites, tags and other relevant sequence elements.

ID	Name	Sequence (5' to 3')
<i>IWR1.Sc full-length for expression</i>		
EC-P005	IWR1-Sc-F-NdeI	GGGCACGGGCATATGAGCACTATTAGTACCACAAC
EC-P006	IWR1-Sc-F-NdeI-His	GGGCACGGGCATATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCATGAGCACTATTAGTACCACAAC
EC-P008	IWR1-Sc-R-XhoI-Stop	GGGCACGGGCTCGAGTTAGCTTCTATTAATTTTTTTCTGTAG
EC-P009	IWR1-Sc-R-XhoI-Stop-His	GGGCACGGGCTCGAGTCACTGTTGGTGGTGGTGGTGGCTTCTATTAATTTTTTTCTG
EC-P011	IWR1-Sc-F-exon1	ATGAGCACTATTAGTACCACAACGG
EC-P012	IWR1-Sc-R-exon1/2	CTCTTTCCCTCATCAATTAATAATGCCTGCACAGAATCC
EC-P013	IWR1-Sc-F-exon1/2	GGATTCTGTGACGATTAATTAATGATGAGGGGAAAAGAG
EC-P014	IWR1-Sc-R-exon2	TTAGCTTCTATTAATTTTTTTCTGTAGC

ID	Name	Sequence (5' to 3')
EC-P181	TFG2-Sc-F-sequencing	CTTCCCTTGCGGAAATG
EC-P182	TFG2-Sc-R-sequencing	TGCATCTAAATACATATGTACAGATAGC
EC-P183	TFG2-Sc-F-SacI	GAAGGGGAGCTTCTCCTTCAGTAGAACCAATTTGTG
EC-P184	TFG2-Sc-R-BamHI	GAAGGGGGATCCCCCATATGGACAACAACG
EC-P185	TFG2-Sc-mutagenesis-PY203AA	ACAGATAGGGATGGTAGAGATAGATATATAGCAGCTGTGAAGACGATTCCEAAAAAAC
EC-P186	TFG2-Sc-mutagenesis-NLSΔ	GGCAAGCTTACCTGAAGAAG-GTGAAGACGATTCCEAAAAAAC
EC-P187	TFG2-Sc-mutagenesis-sequencing	CAACTTGCATGGCCAAGAAC
	<i>Pol I, II, III genomic EGFP-tags</i>	
EC-P188	RPA190-Sc-F-homologous	ATGTTGGTACGGGTTCAATTTGATGTGTTAGCAAAGTTCCAAATGCGGCTCGTACGCTGCAG GTGCGAC
EC-P189	RPA190-Sc-R-homologous	ATAAACTAATATTTAAATCGTAATAATTATGGGACCTTTTGCCTGCTTCTAATCGATGAATTC GAGCTCG
EC-P190	RPA190-Sc-F-sequencing	TGATGACCAGGCAAGGTAC
EC-P191	RPA190-Sc-R-sequencing	CCGCTTGATGATGATGCAC
EC-P192	RPB1-Sc-F-homologous	ATTCTCCAAAGCAAGCAACAAAAGCATAATGAAAATGAAAATCCAGACGTACGCTGCA GGTTCGAC
EC-P193	RPB1-Sc-R-homologous	CTATATATAATGTAATAACGTCAAATACGTAAGGATGATATACTATATCAATCGATGAATT CGAGCTCG
EC-P194	RPB1-Sc-F-sequencing	CAGCCCAACGTCTCCAG
EC-P195	RPB1-Sc-R-sequencing	GCATTTACTAGCGCCGTTG
EC-P196	C160-Sc-F-homologous	AGCGATGTCTATTTGAAAGTCTCTCAAATGAGGCAGCTTTAAAAGCGAACCGTACGCTGCAG GTGCGAC
EC-P197	C160-Sc-R-homologous	GTAGAAAAATAATACAAATGCTATAAAAAAGTTAAAAACGACTACTTTAATCGATGAATT CGAGCTCG
EC-P198	C160-Sc-F-sequencing	GCAGCTAGCCTCCTTGG
EC-P199	C160-Sc-R-sequencing	AGCTTGTGAGTGCATACCAG
EC-P200	RPB4-Sc-F-homologous	ATGAGTTGGAAAGGATACTAAAGGAATTGTCAAACCTAGAAACACTCTATCGTACGCTGCA GGTTCGAC
EC-P201	RPB4-Sc-R-homologous	CAGGGTGAAAATTCATTACTAGTTACACACACGTATACATACAGTTAATCGATGAATT CGAGCTCG
EC-P202	RPB4-Sc-F-sequencing	CAAATTTCTCCGATTTAGAGACC
EC-P203	RPB4-Sc-R-sequencing	TCGATCAGAGGTGTAGATGC
	<i>RPB4.Sc deletion</i>	
EC-P204	RPB4-Sc-deletion-natMX4-F	TACACAAATATTCATCCGTATTGAAAATATATATATAAGAATATAGAAAACGTACGCTGCA GGTTCGAC
EC-P205	RPB4-Sc-deletion-natMX4T-R	CCACAGGGTGAAAATTCATTACTAGTTACACACACGTATACATACAGTTAATCGATGAATTC GAGCTCG
EC-P206	RPB4-Sc-F-sequencing	CCAGAAGGGTAAATCTGAGG
EC-P207	RPB4-Sc-R-sequencing	GGCCAGCGAAGAGACC

2.1.4 Plasmids

Plasmids (Table 5) were generated by classical cloning, yeast *in vivo* recombination or site-directed mutagenesis (see 2.2.2). Inserts were amplified in single- or multi-step PCRs using the indicated oligonucleotides. Plasmid identifiers (IDs) are provided for both the original Elmar Czeko plasmid collection (pEC) and for the Patrick Cramer laboratory plasmid collection (pPC).

Table 5 · Plasmids

ID	Plasmid	Type	Cloning			
			from	into	using oligonucleotides	via
<i>IWR1.Sc full-length expression</i>						
pEC003 pPC1317	IWR1.Sc-His6	pET-21b	genomic DNA	pET-21b	EC-P024, EC-P036	classical cloning (NdeI, XhoI)
pEC005 pPC1318	His6-IWR1.Sc	pET-21b	pEC003	pET-21b	EC-P006, EC-P035	classical cloning (NdeI, XhoI)
pEC023 pPC1319	IWR1.Sc-Strep	pET-21b	pEC003	pET-21b	EC-P024, EC-P038	classical cloning (NdeI, XhoI)
<i>IWR1.Sc minimal Pol II binding</i>						
pEC040 pPC1320	His6-IWR1.Sc [1-218]	pET-21b	pEC003	pET-21b	EC-P062, EC-P065	classical cloning (NdeI, XhoI)
pEC041 pPC1321	His6-IWR1.Sc [1-230]	pET-21b	pEC003	pET-21b	EC-P062, EC-P066	classical cloning (NdeI, XhoI)
pEC043 pPC1322	His6-IWR1.Sc [9-245]	pET-21b	pEC003	pET-21b	EC-P063, EC-P068	classical cloning (NdeI, XhoI)
pEC044 pPC1323	His6-IWR1.Sc [19-245]	pET-21b	pEC003	pET-21b	EC-P064, EC-P068	classical cloning (NdeI, XhoI)
<i>IWR1.Hs synthetic gene</i>						
pEC047 pPC1324	IWR1.Hs	pMA	Mr. Gene	pMA	-	chemical synthesis
<i>IWR1.Sc refined minimal Pol II binding</i>						
pEC058 pPC1325	His6-IWR1.Sc [9-201]	pET-21b	pEC003	pET-21b	EC-P063, EC-P086	classical cloning (NdeI, XhoI)
pEC059 pPC1326	His6-IWR1.Sc [9-211]	pET-21b	pEC003	pET-21b	EC-P063, EC-P087	classical cloning (NdeI, XhoI)
pEC061 pPC1327	His6-IWR1.Sc [9-223]	pET-21b	pEC003	pET-21b	EC-P063, EC-P088	classical cloning (NdeI, XhoI)
<i>IWR1.Sc for in vivo analysis</i>						
pEC069 pPC1328	IWR1.Sc [promoter+3'UTR]	pRS316	genomic DNA	pRS316	EC-P097, EC-P104	classical cloning (XbaI, XhoI)
pEC074 pPC1329	IWR1.Hs [promoter+3'UTR from Sc]	pRS317	pEC047	pEC069	EC-P133, EC-P134	yeast <i>in vivo</i> recombination (BglII, BstEII)
pEC076 pPC1330	IWR1.Sc (1-353)	pRS318	pEC003	pEC074	EC-P135, EC-P141	yeast <i>in vivo</i> recombination (EcoRI)
pEC077 pPC1331	IWR1.Sc (9-223)	pRS319	pEC003	pEC074	EC-P136, EC-P140	yeast <i>in vivo</i> recombination (EcoRI)
pEC078 pPC1332	IWR1.Sc (55-353)	pRS320	pEC003	pEC074	EC-P137, EC-P141	yeast <i>in vivo</i> recombination (EcoRI)
pEC079 pPC1333	IWR1.Sc (114-353)	pRS321	pEC003	pEC074	EC-P138, EC-P141	yeast <i>in vivo</i> recombination (EcoRI)
pEC080 pPC1334	IWR1.Sc (224-353)	pRS322	pEC003	pEC074	EC-P139, EC-P141	yeast <i>in vivo</i> recombination (EcoRI)
pEC081 pPC1335	IWR1.Sc (1-353, dNES)	pRS316	pEC094	pEC074	EC-P135, EC-P141	yeast <i>in vivo</i> recombination (EcoRI)
pEC082 pPC1336	IWR1.Sc (9-223, dNES)	pRS316	pEC094	pEC074	EC-P136, EC-P140	yeast <i>in vivo</i> recombination (EcoRI)
pEC083 pPC1337	IWR1.Sc (55-353, dNES)	pRS316	pEC094	pEC074	EC-P137, EC-P141	yeast <i>in vivo</i> recombination (EcoRI)
pEC084 pPC1338	IWR1.Sc (114-353, dNES)	pRS316	pEC094	pEC074	EC-P138, EC-P141	yeast <i>in vivo</i> recombination (EcoRI)
pEC085 pPC1339	IWR1.Sc (1-353)-EGFP	pRS316	EGFP from pYM28	pEC076	EC-P143, EC-P144	yeast <i>in vivo</i> recombination (BamHI)
pEC086 pPC1340	IWR1.Sc (9-223)-EGFP	pRS316	EGFP from pYM28	pEC077	EC-P142, EC-P144	yeast <i>in vivo</i> recombination (BamHI)

ID	Plasmid	Type	Cloning			
pEC087 pPC1341	IWR1.Sc (55-353)- EGFP	pRS316	EGFP from pYM28	pEC078	EC-P143, EC-P144	yeast <i>in vivo</i> recombination (BamHI)
pEC088 pPC1342	IWR1.Sc (114-353)- EGFP	pRS316	EGFP from pYM28	pEC079	EC-P143, EC-P144	yeast <i>in vivo</i> recombination (BamHI)
pEC089 pPC1343	IWR1.Sc (224-353)- EGFP	pRS316	EGFP from pYM28	pEC080	EC-P143, EC-P144	yeast <i>in vivo</i> recombination (BamHI)
pEC090 pPC1344	IWR1.Sc (1-353, dNES)-EGFP	pRS316	EGFP from pYM28	pEC081	EC-P143, EC-P144	yeast <i>in vivo</i> recombination (BamHI)
pEC091 pPC1345	IWR1.Sc (9-223, dNES)-EGFP	pRS316	EGFP from pYM28	pEC082	EC-P142, EC-P144	yeast <i>in vivo</i> recombination (BamHI)
<i>IWR1.Sc NES deletion and NLS mutants</i>						
pEC094 pPC1346	His6-IWR1(d182-190, dNES).Sc	pET-21b	pEC005	-	EC-P145	site-directed mutagenesis
pEC095 pPC1347	His6-IWR1(RVKRR14- 18AVAAA).Sc	pET-21b	pEC005	pET-21b	EC-P006, EC-P153, EC-P154, EC-P035	2-step PCR mutagenesis, classical cloning (NdeI, XhoI)
pEC096 pPC1348	His6-IWR1(KRVK33- 36AAVA).Sc	pET-21b	pEC005	pET-21b	EC-P006, EC-P155, EC-P156, EC-P035	2-step PCR mutagenesis, classical cloning (NdeI, XhoI)
pEC097 pPC1349	His6-IWR1(14-18/33- 36 mutated).Sc	pET-21b	pEC095	pET-21b	EC-P006, EC-P153, EC-P154, EC-P035	2-step PCR mutagenesis, classical cloning (NdeI, XhoI)
pEC098 pPC1350	IWR1(RVKRR14- 18AVAAA).Sc	pRS316	pEC076	-	EC-P097, EC-P153, EC-P154, EC-P104	2-step PCR mutagenesis, classical cloning (XbaI, XhoI)
pEC099 pPC1351	IWR1(KRVK33- 36AAVA).Sc	pRS316	pEC076	-	EC-P097, EC-P155, EC-P156, EC-P104	2-step PCR mutagenesis, classical cloning (XbaI, XhoI)
pEC100 pPC1352	IWR1(14-18/33-36 mutated).Sc	pRS316	pEC098	-	EC-P097, EC-P153, EC-P154, EC-P104	2-step PCR mutagenesis, classical cloning (XbaI, XhoI)
pEC101 pPC1353	IWR1(RVKRR14- 18AVAAA)-EGFP.Sc	pRS316	EGFP from pYM28	pEC098	EC-P143, EC-P144	yeast <i>in vivo</i> recombination (BamHI)
pEC102 pPC1354	IWR1(KRVK33- 36AAVA)-EGFP.Sc	pRS316	EGFP from pYM28	pEC099	EC-P143, EC-P144	yeast <i>in vivo</i> recombination (BamHI)
pEC103 pPC1355	IWR1(14-18/33-36 mutated)-EGFP.Sc	pRS316	EGFP from pYM28	pEC100	EC-P143, EC-P144	yeast <i>in vivo</i> recombination (BamHI)
pEC104 pPC1356	IWR1(RVKRR14- 18AVAAA, dNES)- EGFP.Sc	pRS316	pEC101	-	EC-P145	site-directed mutagenesis
pEC105 pPC1357	IWR1(KRVK33- 36AAVA, dNES)- EGFP.Sc	pRS316	pEC102	-	EC-P145	site-directed mutagenesis
pEC106 pPC1358	IWR1(14-18/33-36 mutated, dNES)- EGFP.Sc	pRS316	pEC103	-	EC-P145	site-directed mutagenesis
<i>TFG2.Sc rescue plasmids</i>						
pEC119 pPC1359	TFG2 rescue in pRS316	pRS315	genomic DNA	pRS315	EC-P183, EC-P184	classical cloning (SacI, BamHI)
pEC120 pPC1360	TFG2 rescue in pRS315	pRS316	genomic DNA	pRS316	EC-P183, EC-P184	classical cloning (SacI, BamHI)
pEC121 pPC1361	TFG2 (PY203/204AA) rescue in pRS315	pRS316	pEC120	-	EC-P185	site-directed mutagenesis
pEC122 pPC1362	TFG2 (d165-204) rescue in pRS315	pRS316	pEC120	-	EC-P186	site-directed mutagenesis

2.1.5 Reagents and Consumables

Chemicals were obtained from Merck (Darmstadt, Germany), Roth (Karlsruhe, Germany) or Sigma-Aldrich (Seelze, Germany), unless otherwise stated. Cloning enzymes and reagents were obtained from Fermentas (St. Leon-Rot, Germany), New England Biolabs (Frankfurt am Main, Germany) or Agilent/Stratagene (Waldbronn, Germany). For DNA preparation commercial kits from Qiagen (Hilden, Germany) were used. DNA oligonucleotides were ordered at ThermoScientific (Ulm, Germany). Antibodies were generally supplied by abcam (Cambridge, UK) or Santa Cruz Biotechnology (Heidelberg, Germany). Reagents and consumables were ordered at standard laboratory suppliers.

2.1.6 Media and additives

Media were usually taken from lab stocks (Table 6). Media additives were sterile filtered (Table 7).

Table 6 · Media for *Escherichia coli* and *Saccharomyces cerevisiae*

Media	Application	Description
LB	<i>Escherichia coli</i> culture	1% (w/v) tryptone; 0.5% (w/v) yeast extract; 0.5% (w/v) NaCl
SOB	<i>Escherichia coli</i> transformation	2% (w/v) tryptone; 0.5% (w/v) yeast extract; 8.55 mM NaCl; 2.5 mM KCl; 10 mM MgCl ₂
SOC	<i>Escherichia coli</i> transformation	SOB + 20 mM glucose (before use)
YPD	Yeast culture	2% (w/v) peptone; 2% (w/v) glucose; 1% (w/v) yeast extract
Synthetic complete (SC) amino acid drop-out medium	Yeast culture	0.69% (w/v) nitrogen base; amino acid drop out mix (w/v) as indicated by manufacturer; 2% (w/v) glucose; pH 5.6-6.0
YPD plates	Yeast plate	YPD, 2% (w/v) agar
SC amino acid drop-out plates	Yeast plate	SC amino acid drop-out medium; 2% (w/v) agar
5-FOA plates	Yeast plate	SC (-ura) + 0.01% (w/v) uracil; 0.2% (w/v) 5-FOA; 2% (w/v) agar
Pre-sporulation plates	Yeast plate	1% (w/v) KCH ₃ COO; 0.1% (w/v) yeast extract; 0.079% (w/v) CSM amino acid complete mix; 0.25% (w/v) glucose; pH 5.6-6.0
Sporulation t plate	Yeast plate	1% (w/v) KCH ₃ COO; 0.079% (w/v) CSM amino acid complete mix; pH 5.6-6.0

Table 7 · Media additives for *Escherichia coli* and *Saccharomyces cerevisiae*

Additive	Description	Stock solution	Applied concentration
IPTG	<i>Escherichia coli</i> induction	1 M in H ₂ O	0.5 mM
Ampicillin	Antibiotic	100 mg/ml in H ₂ O	100 µg/ml for <i>Escherichia coli</i> culture, 50 µg/ml for yeast culture
Kanamycin	Antibiotic	30 mg/ml in H ₂ O	30 µg/ml for <i>Escherichia coli</i> culture
Chloramphenicol	Antibiotic	50 mg/ml in EtOH	50 µg/ml for <i>Escherichia coli</i> culture
Tetracyclin	Antibiotic	12.5 mg/ml in 70% EtOH	12.5 µg/ml for yeast culture
Geneticin (G418)	Antibiotic	200 mg/ml in H ₂ O	200 µg/ml for yeast culture
Nourseothricin (cloneNAT)	Antibiotic	100 mg/ml in H ₂ O	100 µg/ml for yeast culture

2.1.7 Buffers and solutions

General buffers, dyes and solutions are listed in Table 8. Purification buffers for HPLC systems were filtered and degassed before usage (Table 9, Table 10, Table 11).

Table 8 · General buffers, dyes and solutions

Name	Application	Description
1x Bradford dye	Determination of protein concentration	1:5 dilution of Bradford concentrate (BioRad)
4x stacking gel buffer	SDS-PAGE	0.5M Tris, 0.4% (w/v) SDS, pH 6.8 at 25°C
4x separation gel buffer	SDS-PAGE	3 M Tris, 0.4% (w/v) SDS, pH 8.9 at 25°C
electrophoresis buffer	SDS-PAGE	25 mM Tris, 0.1% (w/v) SDS, 250 mM glycine
5x SDS sample buffer	SDS-PAGE	250 mM Tris/HCl pH 7.0 at 25°C, 50% (v/v) glycerol, 0.5% (w/v) bromophenol blue, 7.5% (w/v) SDS, 12.5% (w/v) β -mercaptoethanol
Gel staining solution	Coomassie staining	50% (v/v) ethanol, 7% (v/v) acetic acid, 0.125% (w/v) Coomassie Brilliant Blue R-250
Gel destaining solution	Coomassie staining	5% (v/v) ethanol, 7.5% (v/v) acetic acid
1x TBE	Agarose gel electrophoresis	8.9 mM Tris, 8.9 mM boric acid, 2 mM EDTA (pH 8.0, 25°C)
1x TE	Nucleic acids, yeast transformation	10 mM Tris pH 7.4, 1mM EDTA
10x TBS	Agarose gel electrophoresis	500 mM Tris/HCl pH 7.5, 1.5 M NaCl
1x PBS	Divers	2 mM KH ₂ PO ₄ , 4 mM Na ₂ HPO ₄ , 140 mM NaCl, 3 mM KCl, pH 7.4 @ 25°C
6x Loading dye	Agarose gel electrophoresis	1.5 g/l bromophenol blue, 1.5 g/l xylene cyanol, 50% (v/v) glycerol
2x Tris/Glycine transfer buffer	Western Blotting	2.4% (w/v) glycin, 0.8% (w/v) Tris, 40% (v/v) ethanol
1x CAPS transfer buffer	Western Blotting for Edman sequencing	0.221% (w/v) CAPS, 0.05% (w/v) DTT, 15% (v/v) methanol, pH 10.5 (adjusted with NaOH)
pre-treatment solution	lysis for quantitative western blotting	7.5% (v/v) β -mercaptoethanol, 1.85 M NaOH
TFB-1	Chemically competent cells	30 mM KOAc, 50 mM MnCl ₂ , 100 mM RbCl, 10 mM CaCl ₂ , 15% (v/v) glycerol, pH 5.8 at 25°C
TFB-2	Chemically competent cells	10 mM MOPS pH 7.0 at 25°C, 10 mM RbCl, 75 mM CaCl ₂ , 15% (v/v) glycerol

Name	Application	Description
100x PI	Protease inhibitor mix	0.028 mg/ml Leupeptin, 0.137 mg/ml Pepstatin A, 0.017 mg/ml PMSF, 0.33 mg/ml benzamidine, in 100% EtOH p.a.
Avidin	Strep purification	50 µmol/l avidin, 50% (w/v) glycerin, 20 mM Tris pH 8.0, 150 mM NaCl, 10 mM β-mercaptoethanol
50x d-desthiobiotin	Strep-tag purification	125 mM d-desthiobiotin, in 500 mM Tris pH 8.0
50% (w/v) PEG 3,350	Yeast genomic transformation	50% (w/v) PEG 3,350, autoclaved
1 M lithium acetate	Yeast genomic transformation	1 M lithium acetate, pH 7.5 adjusted with acetic acid
Salmon sperm DNA	Yeast genomic transformation	2 mg/ml Salmon sperm DNA
One-step buffer	Yeast plasmid transformation	0.2 M lithium acetate, 40% (w/v) PEG 3,350, 100 mM DTT
Paraformaldehyde fixing	Light microscopy	10% (w/v) paraformaldehyde, 13 mM NaOH, 1 M K ₃ PO ₄

Table 9 · Iwr1 purification buffers

Name	Description
IMAC lysis buffer	20 mM Tris/HCl, pH 8.0, 150 mM NaCl, 10 mM imidazole, 10 β-mercaptoethanol, 1x PI
IMAC washing buffer	20 mM Tris/HCl, pH 8.0, 150 mM NaCl, 20/30/40 mM imidazole, 10 β-mercaptoethanol
IMAC elution buffer	20 mM Tris/HCl pH 8.0, 150 mM NaCl, 250 mM imidazole, 10 β-mercaptoethanol
Strep lysis buffer	20 mM Tris/HCl pH 8.0, 150 mM NaCl, 10 β-mercaptoethanol, 1x PI, 250 pmol avidin per 1 l culture
Strep washing buffer	20 mM Tris/HCl pH 8.0, 150 mM NaCl, 10 β-mercaptoethanol
Strep elution buffer	20 mM Tris/HCl pH 8.0, 150 mM NaCl, 10 β-mercaptoethanol, 2.5 mM desthiobiotin
anion exchange buffer A	20 mM Tris/HCl pH 8.0, 50 mM NaCl, 10 mM DTT
anion exchange buffer B	20 mM Tris/HCl pH 8.0, 2 M NaCl, 10 mM DTT

Table 10 · Pol II purification buffers

Name	Description
3x freezing buffer (used as 4x)	150 mM Tris/HCl (pH 7.9 at 4 °C), 3 mM EDTA, 30 µM ZnCl ₂ , 30% (v/v) glycerol, 3% (v/v) DMSO, 30 mM DTT, 3x PI
HSB 150	50 mM Tris/HCl (pH 7.9 at 4 °C), 150 mM KCl, 1 mM EDTA, 10 µM ZnCl ₂ , 10% (v/v) glycerol, 10 mM DTT, 1x PI
HSB 0/0	50 mM Tris/HCl (pH 7.9 at 4 °C), 1 mM EDTA, 10 µM ZnCl ₂ , 10% (v/v) glycerol, 2.5 mM DTT, 1x PI
HSB 1000/0	50 mM Tris/HCl (pH 7.9 at 4 °C), 1000 mM KCl, 1 mM EDTA, 10 µM ZnCl ₂ , 10% (v/v) glycerol, 2.5 mM DTT, 1x PI
HSB 1000/7	50 mM Tris/HCl (pH 7.9 at 4 °C), 1000 mM KCl, 7 mM imidazole, 1 mM EDTA, 10 µM ZnCl ₂ , 10% (v/v) glycerol, 2.5 mM DTT, 1x PI
5x IMAC buffer	100 mM Tris/HCl (pH 7.9 at 4 °C), 750 mM KCl, 50 µM ZnCl ₂
Ni buffer 7, 50, 100	20 mM Tris/HCl (pH 7.9 at 4 °C), 150 mM KCl, 7/50/100 mM imidazole, 10 µM ZnCl ₂ , 2.5 mM DTT, 1x PI
5x anion exchange buffer	100 mM Tris-acetate (pH 7.9 at 4 °C), 50% (v/v) glycerol, 2.5 mM EDTA pH 7.9, 50 µM ZnCl ₂
MonoQ 0	20 mM Tris-acetate (pH 7.9 at 4 °C), 10% (v/v) glycerol, 0.5 mM EDTA pH 7.9, 10 µM ZnCl ₂ , 10 mM DTT
MonoQ 150	20 mM Tris-acetate (pH 7.9 at 4 °C), 150 mM KOAc, 10% (v/v) glycerol, 0.5 mM EDTA pH 7.9, 10 µM ZnCl ₂ , 10 mM DTT
MonoQ 2000	20 mM Tris-acetate (pH 7.9 at 4 °C), 2000 mM KOAc, 10% (v/v) glycerol, 0.5 mM EDTA pH 7.9, 10 µM ZnCl ₂ , 10 mM DTT
Pol II size exclusion buffer	5 mM HEPES (pH 7.25 at 20 °C), 40 mM (NH ₄) ₂ SO ₄ , 10 µM ZnCl ₂ , 10 mM DTT

Table 11 · Rpb4/7 purification buffers

Name	Description
Rpb4/7 freezing buffer	50 mM Tris (pH 7.0 at 4 °C), 150 mM NaCl, 10% glycerol, 10 mM β -mercaptoethanol, 1 \times PI
IMAC buffer	50 mM Tris (pH 7.5 at 4 °C), 150 mM NaCl, 0/10/20/50/200 mM imidazole, 10 mM β -mercaptoethanol, 1 \times PI
IMAC salt buffer	50 mM Tris (pH 7.5 at 4 °C), 2 M NaCl, 10 mM β -mercaptoethanol, 1 \times PI
SourceQ 0	20 mM Tris (pH 7.5 at 4 °C), 1 mM EDTA, 10 mM DTT
SourceQ 100	20 mM Tris (pH 7.5 at 4 °C), 100 mM NaCl, 1 mM EDTA, 10 mM DTT
SourceQ 2000	20 mM Tris (pH 7.5 at 4 °C), 2 M NaCl, 1 mM EDTA, 10 mM DTT
Pol II buffer	5 mM HEPES (pH 7.25 at 20 °C), 40 mM (NH ₄) ₂ SO ₄ , 10 μ M ZnCl ₂ , 10 mM DTT

2.2 General methods

2.2.1 Preparation and transformation of competent bacterial cells

Chemically competent *Escherichia coli* cells were prepared from a 200 ml LB culture containing appropriate antibiotics. After inoculation at an OD₆₀₀ of 0.05 with an overnight pre-culture, cells were grown at 37 °C to an OD₆₀₀ of 0.5 and chilled on ice for 10 min. For all following steps, care was taken for the samples to be constantly kept at 4 °C. Cells were centrifuged at 3,200 g for 10 min, thoroughly washed with 50 ml TFB-1 buffer, pelleted again and thoroughly resuspended in 4 ml TFB-2 buffer. 50 μ l aliquots were frozen in liquid nitrogen and cells stored at -80 °C. For transformation, 100 ng of a plasmid or 5 μ l from a ligation reaction was added to an aliquot of competent cells, followed by incubation on ice for 20 min and a 45 s heat shock at 42 °C in a water bath with subsequent cooling on ice for 1 min. After addition of 250 μ l LB medium the transformation mix was incubated in a thermomixer shaking at 1,400 rpm for 1 h at 37 °C. Finally, cells were plated onto selective LB plates and incubated overnight at 37 °C.

2.2.2 Cloning and mutagenesis

As *Escherichia coli* expression vectors the Novagen pET series and as *Saccharomyces cerevisiae* vectors the pRS series obtained from Euroscarf was used.

For polymerase chain reaction (PCR) primers the following properties were considered: 1) A 6 bp GC-rich sequence was added at the primer's 5' end next to a restriction sites; 2) complementary regions consisted of at least 18 bp, the melting temperature was targeted to be at 55 °C as calculated with ApE (Wayne Davis, 2009) and the region concluded with a G or C nucleotide; 3) hairpins and complementary regions between primer pairs were avoided. Purification tags were either introduced via the primer or by in-frame cloning in the vector. PCR reactions were carried out with Herculase or Herculase II polymerases (Stratagene), Phusion polymerase (Finnzymes), Pfu polymerase (Fermentas) or Taq polymerase (Fermentas), depending on requirements. Reaction were typically performed at 50 µl scale and contained polymerase specific buffers, 1-30 ng of plasmid DNA or 100-200 ng of genomic DNA template, 0.2 mM dNTP-mix, 0.5 µM forward and reverse primer each, variable amounts of polymerase (0.5 to 5.0 U), and optionally MgCl₂ or DMSO. For test PCRs the reaction volume was usually scaled down to 20 µl. PCR programs were optimized for each reaction, in particular in terms of annealing temperature (using a gradient PCR cycler) and elongation times.

PCR products were purified using PCR-purification kits (Qiagen or Metabion). PCR products and vectors were digested with restriction endonucleases (NEB or Fermentas) under conditions recommended by the supplier. Digested vectors were dephosphorylated by addition of 1.0 U FastAP enzyme (Fermentas) and incubation for 10 min at 37 °C followed by heat inactivation for 5 min at 75 °C. Samples were separated by agarose gel electrophoresis (typically 1% w/v) in 1× TBE buffer and visualized by ethidiumbromide or SYBR Safe DNA staining (1:10,000, Invitrogen). DNA fragments were extracted and purified using the QIAquick gel extraction kit (Qiagen). PCR products and linearized vectors were ligated using 5.0 U of T4

DNA ligase (Fermentas) in a 20 μ l reaction volume for 1 h at 20 °C or overnight at 18 °C. Usually a molar vector to insert ratio of 1:2 was applied, calculated based on the DNA concentration and – as a proxy for the molecular weight of the nucleic acids – their length in base pairs. Ligation products were transformed as described (see 2.2.1). 5 ml overnight cultures were inoculated with single clones from selective LB plates. Plasmids were isolated from the *Escherichia coli* clones using miniprep purification kits (Qiagen or Metabion) and verified by restriction analysis and DNA sequencing (MWG).

Site-directed mutagenesis was applied to introduce point mutations into plasmids. Usually, 10 ng of template plasmid was used in the mutational PCR with Pfu polymerase (Fermentas). Primers typically contained at least 15 bp complementary regions on both sides of a mutation. PCR reactions were performed with only one primer and annealing temperatures between 45 °C and 55 °C and 2 min per kb extension time. Next, the template vector was digested with 10 U DpnI (Fermentas) for 1 h at 37 °C and 5 μ l of the reaction were directly used for transformation of competent *Escherichia coli* cells. For the generation of mutations that were not amenable to site-directed mutagenesis a 2-step PCR approach was used. Specifically, two overlapping PCR fragments were generated in a first step that contained the desired mutation introduced via primers. Next, the two PCR fragments were used as a template in a second PCR to amplify a single fragment carrying the mutation, which could be cloned into a vector.

As an alternative to restriction enzyme based classical cloning, *in vivo* homologous recombination in yeast was applied for plasmids carrying both selectable markers in yeast and *Escherichia coli*. In addition to the general robustness of the technique, the location of the insertion can be chosen freely under the single constraint that a restriction site is present on the plasmid only between the two homologous regions used for recombination, but not outside of them. Briefly, yeast cells were transformed as described (see 2.3.2) with 1 μ g of digested and gel-purified plasmid and 2- to 3-fold molar excess of the PCR-generated insert containing 50 bp homologous regions at its 5' and 3' ends. After two to four days up to five clones were picked from the selective plates for overnight growth in 5 ml selective medium at 30 °C with subsequent yeast plasmid preparation according to a modified version of the Qiagen protocol for

DNA preparation from *Escherichia coli* (Jones, 2001). 5 µl of the obtained yeast plasmid DNA was transformed into *Escherichia coli* cells (see 2.2.1) followed by selection of a single clone per original yeast clone, growth of a 5 ml culture at 37 °C, plasmid preparation (Qiagen), test digestion and sequencing (MWG).

2.2.3 Protein expression in *Escherichia coli*

Recombinant proteins were routinely expressed in *Escherichia coli* BL21-CodonPlus (DE3) RIL cells (Stratagene). Plasmids carrying desired protein variants were transformed as described (see 2.2.1). Cells were grown at 37 °C in LB medium with required antibiotics up to an OD₆₀₀ of 0.5 to 0.8. Cultures were then cooled on ice for 30 min, induced by addition of 0.5 mM IPTG with subsequent overnight expression at 18 °C. Cells were harvested by centrifugation at 3,500 g and 4 °C for 20 min, washed in lysis buffer, flash frozen in liquid nitrogen and stored at -80° C.

2.2.4 Protein analysis

Determination of protein concentrations

Total protein concentrations were usually determined with a Bradford assay (Bradford, 1976) by measuring the absorption at 595 nm in 1 ml of a dye reagent solution (Bio-Rad) after addition of typically 1 µl to 5 µl of protein. Reference curves were prepared for each new batch of dye reagent using bovine serum albumin as a standard (Fraktion V, Roth). Alternatively, protein concentrations were calculated via the absorption at 280 nm using a ND-1000 (NanoDrop) spectrophotometer. Molar absorption coefficients were calculated with ProtParam based on the protein sequence (Gasteiger et al, 2005).

SDS-polyacrylamide gel electrophoresis

Electrophoretic separation of proteins was routinely conducted by SDS-PAGE in BioRad gel systems with 15% acrylamide/0.4% bisacrylamide gels (Laemmli, 1970). For sharper resolution, commercial NuPAGE Novex Midi Bis-Tris 4-12% gradient gels (Invitrogen) were used according to the manufacturer's instructions. Gels were stained with Coomassie solution for 20 min and destained overnight in gel destaining solution (Table 8).

Limited proteolysis analyses

Limited proteolysis time courses were performed to identify stable and compactly folded protein regions. Protein was digested with chymotrypsin or trypsin protease at 37 °C and samples were taken at 30 s and 1, 3, 8, 15 and 30 min. Digests were performed with 20 µg protein and 0.1 µg protease per time point in the Iwr1 size exclusion buffer supplemented with 5 mM CaCl₂. The reactions were stopped by the addition of 1× SDS sample buffer and immediate heating at 95 °C for 5 min. After separation of degraded protein fragments by SDS-PAGE bands were excised and analyzed by Edman sequencing (Niall, 1973).

Edman sequencing

For N-terminal sequencing, proteins were separated by a SDS-PAGE, stained with Coomassie blue and transferred either by western blotting or by passive adsorption onto PVDF membranes. Western blotting was performed in CAPS transfer buffer at 100 V for 1 h or 20 V overnight at 4 °C. For passive adsorption, bands of interest were excised from the gel, dried in a speed-vac and rehydrated in 20 µl swelling buffer at room temperature. Next, 100 µl of double-distilled water and a piece of PVDF membrane, previously soaked in ethanol, were added (Schleicher & Schuell BioScience). Once the solution turned blue, 10 µl of methanol was added and the sample was incubated for up to 4 days until the transfer was complete, as indicated by the migration of the blue colour from the solution onto the PVDF membrane. The membrane was washed five times by vortexing for 30 s in 10% methanol. Finally, the membrane-bound protein was N-terminally sequenced in a PROCISE 491 sequencer (Applied Biosystems) based on automated Edman degradation (Niall, 1973).

2.2.5 Quantitative western blots

To quantify the relative amounts of Rpb1 in the wild-type and $\Delta iwr1$ strains an alkaline lysis of yeast cells was performed. Briefly, 5 ml of yeast grown in exponential phase to an OD_{600} of 1.0 were washed, resuspended in 150 μ l pre-treatment solution, vortexed and incubated on ice for 20 min. Next, 150 μ l of 55% TCA were added followed by another 20 min incubation on ice with intermittent vortexing. Cells were pelleted, resuspended in 85 μ l of 1 \times sample buffer and boiled at 95 °C for a short period of time (90 s) to avoid degradation of large proteins. After centrifugation the absorption at 280 nm of a 1:200 dilution of each sample was measured, using a 1:200 dilution of 1 \times sample buffer as a blank. If required, samples were diluted to the same absorbance and an equivalent of 0.2 absorbance units (as referred to the measurement of the 1:200 dilution) was diluted to 10 μ l and loaded onto an SDS gel. Then the gel was blotted onto a nitrocellulose membrane for 1 h at 100 V in the cold room and the membrane was stained with Ponceau afterwards to confirm an even transfer of protein. The membrane was blocked in 1 \times PBS with 2% (w/v) milk powder for 1 h at room temperature and cut into two parts approximately at the height of a 70 kDa marker band. All following antibody incubations were performed in 10 ml 1 \times PBS with 2% (w/v) milk powder, while washing was done with four consecutive 5 min incubations in 1 \times PBS. The upper part of the membrane was incubated with 8WG16 antibody against Rpb1 (Covance, 1:500, mouse) and the lower part in anti-Pgk1 (Invitrogen, 1:20,000, mouse) for 1 h at room temperature. After washing, both membranes were incubated together in anti-mouse antibody coupled to horse-reddish-peroxidase (BioRad, 1:2,000, goat) for 1 h at room temperature, followed by washing and detection of the transferred protein with the ECL Plus Western Blotting Detection Reagent (GE Healthcare) on a LAS-3000 system (Fujifilm, standard sensitivity, normal binning, 1 min increments). Exposure times were up to 10 min, images were saved as 16 bit TIFF files and quantification was done with ImageJ (rsbweb.nih.gov/ij/). Rpb1 band intensity was normalized by the Pgk1 band intensity and mean values of the ratios were calculated from three independent lysis samples. The linear range of the antibodies was confirmed by 1:2 step serial dilutions.

2.2.6 Bioinformatic tools

Protein and gene sequences were retrieved from the NCBI (www.ncbi.nlm.nih.gov) or the SGD (www.yeastgenome.org) database. Sequence data was displayed and edited using VectorNTI (Invitrogen) and ApE (biologylabs.utah.edu/jorgensen/wayned/ape). Bioinformatic analysis were performed using the Bioinformatics Toolkit (Biegert et al, 2006). Multiple sequence alignments were generated using MUSCLE (Edgar, 2004) and displayed with ESPript (Gouet et al, 1999). Sequence conservation was calculated with EMBOSS (www.ebi.ac.uk) (Needleman & Wunsch, 1970). Protein secondary structures were predicted by HHpred (Soding et al, 2005). NLS sequences were identified and scored with cNLS Mapper (Kosugi et al, 2009b).

2.3 Yeast genetics and assays

2.3.1 Isolation of genomic DNA from yeast

Genomic DNA from yeast was isolated using the DNeasy Blood & Tissue Kit (Qiagen) following the manufacturer's yeast-specific protocol on a QIAcube robot (Qiagen).

2.3.2 Transformation of yeast cells

Chemically competent yeast cells were prepared from a 100 ml culture inoculated with cells from an overnight pre-culture at an OD₆₀₀ of 0.2 and grown to an OD₆₀₀ of 0.8. Cells were centrifuged at 1,250 g for 5 min, washed with sterile water and centrifuged again. The pellets were resuspended in 1 ml 100 mM lithium acetate, split into two 1.5 ml reaction tubes, centrifuged at 18,000 g for 15 s and resuspended in 500 µl per tube. Cells were vortexed, pooled, incubated for 1 h at room temperature or overnight at 4 °C and finally

split into 100 µl aliquots. After centrifugation, supernatants were discarded and 240 µl 50% (w/v) PEG 3,350 was added. If not stored at -80 °C, cells were directly used for transformation at highest competence by adding 36 µl 1 M lithium acetate, 50 µl 2 mg/ml salmon sperm DNA (previously heated at 95 °C for 5 min and afterwards kept on ice) and 1 µg DNA in 34 µl water. Samples were vortexed vigorously for 1 min, incubated at 30 °C for 30 min, followed by 42 °C for 25 min and pelleted at 5,200 g for 15 sec. If required, cells were resuspended in 1 ml YPD and recovered for 3 h at 30 °C to develop antibiotic resistance and then centrifuged at 1,250 g for 5 min. Finally, cells were resuspended in 200 µl 1× TE buffer and plated onto selective plates.

2.3.3 Gene disruption and epitope-tagging

For gene deletion and tagging *Saccharomyces cerevisiae* cells were chemically transformed with the required PCR amplified cassette that was targeted to a certain locus via its 50 bp homologues ends (Janke et al, 2004). After growth of transformants on selective plates, colonies were restreaked and then tested for successful integration of the cassette by PCR with locus- and cassette-specific primers. If necessary, the PCR product was additionally sequenced or test digested.

2.3.4 Long-term storage of yeast strains

For long-term storage, yeast strains were grown on a YPD or selective plate, transferred into 1 ml 30% glycerol in 1.5 ml reaction tubes and then stored at -80 °C. For re-plating, frozen material was scratched from the reaction tube using a glass pipette and then directly transferring it onto a plate.

2.3.5 Determination of yeast doubling times

Doubling times of strains carrying rescue plasmids were based on OD₆₀₀ measurements in -URA medium. Cultures were inoculated at an OD₆₀₀ of 0.1 from an overnight pre-culture. After 2 h the first measurement was taken followed by three more measurements during exponential growth over a period of 7 h or until an OD₆₀₀ of 0.8 was reached to avoid the requirement of diluting the sample for the measurement. Doubling times were calculated as means from three independent cultures.

2.3.6 Determination of protein half-lives

For determination of protein half-lives, culture were grown to an OD₆₀₀ of 0.8. After addition of the translational inhibitor cycloheximide to a final concentration of 300 µg/ml in the culture, samples containing an equivalent of 5 ml of yeast cells at an OD₆₀₀ of 1.0 were collected at each time point over a period of 45 min. Half-lives were calculated by linear regression on the logarithmised relative protein levels determined by quantitative western blot (see 2.2.5).

2.3.7 *In vitro* transcription assays

In vitro transcription assays were performed by Martin Seizl. Nuclear extracts were prepared from 3 l of culture as described by the Hahn laboratory (www.fhcrc.org/labs/hahn/). Plasmid transcription was performed essentially as described (Ranish & Hahn, 1991). Transcription reactions were carried out in a 25 µl volume. The reaction mixture contained 100 µg nuclear extract, 150 ng of pSH515 plasmid, 1× transcription buffer (10 mM HEPES pH 7.6, 50 mM potassium acetate, 0.5 mM EDTA, and 2.5 mM magnesium acetate), 2.5 mM DTT, 192 µg of phosphocreatine, 0.2 mg of creatine phosphokinase, 10 U of RNase

inhibitor (GE Healthcare), and 100 μ M nucleoside triphosphates. For activated transcription, 150 ng of Gal4-VP16 or 200 ng of Gal4-Gal4AH was added. The reaction was incubated at room temperature for 40 min and then stopped with 180 μ l of 100 mM sodium acetate, 10 mM EDTA, 0.5% sodium dodecyl sulphate, and 17 μ g of tRNA/ml. Samples were extracted with phenol-chloroform and precipitated with ethanol. Transcripts were analysed by primer extension essentially as described (Ranish & Hahn, 1991). Instead of the 32 P-labeled lacI oligo, 0.125 pmol of a fluorescently labelled 50-FAM-oligo was used. Quantification was performed with a Typhoon 9400 and the ImageQuant Software (GE Healthcare).

2.4 Fluorescence microscopy

2.4.1 Paraformaldehyde fixation of yeast cells

For fixation, 500 μ l of fresh paraformaldehyde solution was added to 1 ml of Rpb1-EGFP, Rpb3-EGFP, Rpb4-EGFP, Rpa190-EGFP, C160-EGFP and Iwr1-EGFP cells with or without IWR1 deletion and grown to an OD₆₀₀ of 0.8 to 1.0. Samples were incubated for 10 min at room temperature, carefully pelleted in a tabletop centrifuge at 4,000 rpm for 2 min, resuspended in 500 μ l 1 \times PBS and stored at 4 $^{\circ}$ C. For preparation of microscopy samples, 12-well slides were washed in 100% ethanol and incubated for 10 min with 0.02% poly-L-lysine. Depending on the pellet size, 5 to 20 μ l of fixed cells were applied (per well, as for all following indications) and the slide was kept in a humid and dark chamber during all ensuing incubation times. Cells were allowed to settle for 10 min, liquid was aspirated off and 20 μ l 1 \times PBS with 1.0 ng/ μ l DAPI was added for 10 min. After two washes with one drop of 1 \times PBS each, 2.5 μ l 1 \times PBS were added, a cover glass was placed on top and the slide was sealed with nail polish and stored in the dark at 4 $^{\circ}$ C.

2.4.2 Fluorescence microscopy data collection

Imaging was performed with a Leica AF6000 fluorescence microscope at 100× magnification and the Leica LAS AF Software. For DIC images, the DAPI channel and the GFP channel contrast was set to “TL-DIC”, “Fluo” and “Fluo”, respectively and the filter cube was set to “Ana”, “Tri” and “GFP”, respectively. Exposure time and intensity were adjusted as needed to have optimal contrast and not to overexpose the image. In general, exposure times were between 100 and 200 ms for DIC images, around 500 ms for the DAPI channel and 1 to 2 s for the GFP channel. The intensity was set to the lowest or second lowest level (i.e. 1 or 2) for DIC images and the DAPI channel and to the highest or second highest for the GFP channel (i.e. 5 or 4). Images were exported as TIFF files.

2.4.3 Analysis of localization ratios

Quantification of nuclear and cytoplasmic fluorescence was done with ImageJ (rsbweb.nih.gov/ij). The cytoplasm and the nucleus were manually masked using the DAPI image as a guide. The EGFP intensity was integrated and divided by the size of the mask, yielding the average intensity for each of the two compartments. For each analysed cell, the ratio between average nuclear and cytoplasmic fluorescence was calculated. Mean values were calculated for at least 10 cells per strain carrying a certain rescue construct.

2.4.4 Fluorescence microscopy figure preparation

For EGFP images, linear contrast adjustment was used. For DAPI images, non-linear contrast adjustment was used in some cases to facilitate the identification of the nucleus and to reduce background levels. No contrast adjustment was applied to the images before quantification.

2.5 Preparation of the RNA polymerase II-Iwr1 complex

2.5.1 Purification of recombinant Iwr1 and variants

Escherichia coli BL21 (DE3) RIL cells (Stratagene) transformed with plasmids carrying IWR1 full-length and truncation variants were grown in LB medium at 37 °C to an OD₆₀₀ of 0.5. Expression was induced with 0.5 mM IPTG for 16 h at 18 °C. Cells were lysed by sonication in IMAC lysis buffer (Table 9). After centrifugation, the supernatant was loaded onto a 1 ml Ni-NTA agarose column (Qiagen) equilibrated with IMAC lysis buffer. The column was washed three times with 20 CV each of IMAC washing buffer (Table 9) sequentially containing 20, 30, and 40 mM imidazole, and eluted with IMAC elution buffer (Table 9) containing 250 mM imidazole in 0.5 ml fractions whose protein content was assessed by means of the Bradford assay (see 2.2.4). Proteins were further purified on an ÄKTA HPLC system (GE Healthcare) by anion exchange chromatography with a Mono Q 5/50 column (GE Healthcare). The column was equilibrated with anion exchange buffer A (Table 9), and proteins were eluted with a linear gradient of 20 CV from 100 mM to 1 M NaCl in anion exchange buffer B (Table 9). After concentration, the sample was applied to an ÄKTA HPLC system with a Superdex 75 10/300 or Superose 12 10/300 size exclusion column (GE Healthcare) equilibrated with Pol II size exclusion buffer (Table 10). Protein was concentrated to usually 2 mg/ml with an Amicon Ultra-15 centrifugal filter unit with a molecular weight cut-off at 10 kDa (Millipore) and flash-frozen in liquid nitrogen.

2.5.2 Purification of endogenous RNA polymerase II

Cell lysis

Fermented yeast pellets from the Pol II purification strain (PCY ID 068, Table 3) were provided by Stefan Benkert. Two 200 ml cell suspension were thawed in a water bath at a maximum of 30 °C. All following steps were performed on ice or at 4 °C. 1 ml 100× PI (Table 8) was added and the suspension was transferred to the metal chamber of a BeadBeater (BioSpec Products), containing 200 ml soda lime glass beads (0.5 mm diameter, BioSpec Products). Air bubbles were removed stirring gently with a glass rod. The chamber was completely filled up with HSB 150 (Table 10) before assembly with the impeller. The motor was run for 80 min with a duty cycle of 30 s on and 90 s off. To prevent warming, the lysis chamber was covered with a salt-ice mixture, which was regularly replenished. The lysate was separated from the beads by filtering through a mesh funnel and washed with HSB 150, not exceeding a total volume of 1 l. Finally, the lysate was centrifuged (45 min, 13,690 g, 4 °C).

Ultracentrifugation, ammonium sulphate precipitation and redissolving

The lysate was subjected to ultracentrifugation (90 min, 76,220 g, 4 °C) using an SW28 rotor (Beckman-Coulter). The aqueous phase of the supernatant was collected. To precipitate protein, 291 g fine ground ammonium sulphate were added per litre of the elution fraction which was stirred at around 100 rpm overnight at 4 °C. The solution was centrifuged (2× 45 min, 34,200 g, 4 °C) and the precipitate was dissolved in 140 ml HSB 0/0 (Table 10) per 100 g of ammonium sulphate pellet by stirring 1 to 2 h at 4 °C. Conductivity of the protein solution was reduced to the conductivity corresponding to HSB 1000/7 (Table 10) by dilution with HSB 0/0. The sample was centrifuged (10 min, 34,200 g, 4 °C) and 7 mM imidazole were added.

Ni-NTA affinity chromatography

Ni-NTA agarose (Qiagen) was equilibrated with HSB 1000/7 (Table 10) without DTT. The sample was mixed with 16 ml Ni-NTA agarose and incubated for 60 min, stirring at 4 °C. The suspension was split into three aliquots and applied to a gravity flow column. The resin was washed with five CV HSB 1000/7 and then three CV Ni buffer 7 (Table 10). Three CV each of Ni buffer 50 and Ni buffer 100 (Table 10) were applied to the column. Elution fractions were pooled and its conductivity was set corresponding to MonoQ 150 by addition of MonoQ 0 (Table 10).

Anion exchange chromatography

The Ni-NTA affinity elution was centrifuged (20 min, 17,210 g, 4 °C). A Mono Q 10/100 GL column (GE Healthcare) was equilibrated with Mono Q 150 using an ÄKTA HPLC system (GE Healthcare), the sample was loaded and washed with four CV Mono Q 150. For elution a linear salt gradient over twelve CV starting with 100% Mono Q 150 and finishing with 100% Mono Q 2000 was run. Pol II typically eluted at a conductivity of 50–55 mS/cm and the respective fractions were pooled.

Buffer exchange, ammonium sulphate precipitation and storage

After anion exchange chromatography the pooled sample was diluted and buffer was exchanged in an Amicon Ultra-15 concentrator with a molecular weight cut-off at 100 kDa (Millipore) to Pol II size exclusion buffer (Table 10) followed by ammonium sulphate precipitation, freezing of the pellet in liquid nitrogen and storage at -80 °C.

2.5.3 Purification of the recombinant Rpb4/7 subcomplex

Cells from 8l expression culture were pooled and lysed using a homogeniser (EmulsiFlex-C5, Avestin). Lysates were cleared by centrifugation (2× 30 min, 26,890 g, 4 °C). Four times 2 ml Ni-NTA agarose (Qiagen) were

equilibrated with five CV IMAC buffer without imidazole (Table 11). All subsequent steps were carried out at 4 °C. The lysate was loaded twice by gravity flow. The resin was washed with five CV IMAC buffer without imidazole then with three CV each of IMAC salt buffer (Table 11), IMAC buffer with 10 mM and with 20 mM imidazole. Proteins were eluted applying three CV IMAC buffer with 50 mM and six CV with 200 mM imidazole. Ni-NTA elution fractions were pooled and diluted slowly with SourceQ 0 to the conductivity of SourceQ 100 (Table 11). The sample was filtered using a Millex-GP, filter pore size 0.22 µm (Millipore). The following steps were performed using an ÄKTA HPLC system (GE Healthcare). The anion exchange column (Source 15Q 16/10, GE Healthcare) was equilibrated with a mixture of 5% Source Q 2000 and 95% Source Q 0. The pooled Ni-NTA elution fractions were loaded onto the column and washed with two CV. Elution was performed with a linear salt gradient of 100 to 1000 mM NaCl (5% to 50% Source Q 2000) over ten CV. Rpb4/7 eluted at a conductivity of 28 to 31 mS/cm. Rpb4/7 containing fractions from the anion exchange chromatography were pooled and concentrated to a volume of 5 ml using an Amicon Ultra-15 centrifugal filter unit with a molecular weight cut-off at 10 kDa (Millipore). The protein solution was filtered as before and subjected to size exclusion chromatography using a HiLoad 26/60 Superdex 75 pg column (GE Healthcare) with Pol II size exclusion buffer. Fractions of interest were pooled and concentrated as before to a concentration of 6 mg/ml. 40 µl aliquots corresponding to 5 nmol Rpb4/7 were frozen in liquid nitrogen and stored at -80 °C.

2.5.4 Assembly of the RNA polymerase II-Iwr1 complex

To assemble the Pol II-Iwr1 complex, pure Pol II, Rpb4/7 and Iwr1 were mixed in a 1:5:5 stoichiometric ratio. The volume was increased to at least 100 µl by adding Pol II buffer. The sample was incubated in a thermomixer at 20 °C over a period of 20 min. Next, size exclusion chromatography was performed with a Superose 6 column (GE Healthcare) and Pol II buffer. Elution fractions of interest were either pooled and concentrated for structural analysis or TCA precipitated

and run on a SDS-PAGE. The Pol II-Iwr1 complex eluted at around 12 ml. Optionally, static light scattering was applied along with the size exclusion chromatography to measure the molecular weight of the eluted complexes and proteins.

2.5.5 Competition of Iwr1 with nucleic acids and GTFs for RNA polymerase II binding

For competition experiments the Pol II-Iwr1 complex was assembled as described (see 2.5.4). Competing species were either added to the assembly mixture or to a preformed Pol II-Iwr1 complex. Nucleic acids, such as the FC* aptamer (Kettenberger et al, 2006) or a minimal promoter scaffold (Damsma et al, 2007) were added in 2-fold excess relative to Pol II in the assembly mixture. Competition was not dependent on the order of addition, i.e. if nucleic acids were added before or after Iwr1. TFIIB (10-fold excess) and TBP/TFIIB (10-fold excess) together with a closed promoter scaffold (4-fold excess, template strand 5'-AAG CTC AAG TAC TTA CGC CTG GTC ATT ACT AGC TGA CCA TAC TGC ATT TTA TAG GCG CCC GC-3', non-template strand fully complementary) were added to a preformed Pol II-Iwr1 complex. The sample was incubated for 30 min at 20 °C in a thermomixer and analysed by size exclusion chromatography, TCA precipitation of Pol II-Iwr1 fractions and SDS-PAGE. Binding of nucleic acids to polymerase was monitored by an increase in the 260/280 nm absorption ratio.

2.5.6 Reciprocal pull-down experiments

To more closely investigate the interaction of Iwr1 with Pol II, Strep-tagged Iwr1 variants were purified (see 2.5.1) for use in reciprocal pull-down experiments with His-tagged polymerase. The Iwr1 variants were incubated with 20 µg His-tagged Pol II in 5-fold molar excess for 20 min at 20 °C in a 50 µl reaction volume. The mixture was then subjected to either 30 µl Ni-NTA agarose

beads (Qiagen) or 20 μ l Strep agarose beads (IBA BioTAGnology) and binding was allowed for 30 min at 20 °C in a thermomixer shaking at 1,400 rpm. Next, three wash steps were performed with 300 μ l Pol II buffer each, including 10 mM imidazole for Ni-NTA agarose beads. Protein was eluted with 20 μ l Pol II size exclusion buffer including 300 mM imidazole for Ni-NTA agarose beads or 2.5 mM desthiobiotin for Strep agarose beads. Post-binding and elution fractions were analysed by SDS-PAGE. As controls to assess unspecific binding, pull-downs with only either Iwr1 variants or Pol II were used.

2.5.7 Static light scattering analysis

Static light scattering analysis was performed using a Superose 6 size exclusion column (GE Healthcare) combined with a triple detector TDA (Viscotek). Runs were performed with Pol II-Iwr1 assembly mixtures (see 2.5.4) containing typically 100 μ g Pol II. Smaller molecular weight peaks, such as those of unbound Rpb4/7 and Iwr1, were used as internal controls for the measurement. For direct comparison the molecular weight of Pol II was measured in a subsequent run with the same experimental setup. Data analysis was performed with the OmniSEC software (Viscotek).

2.6 Structural analysis of the RNA polymerase II-Iwr1 complex

2.6.1 Negative stain

Assembled Pol II-Iwr1 complexes were concentrated to 0.1 mg/ml and applied to glow-discharged carbon coated holey grids (Quantifoil R3/3, 2 nm carbon on top). Negative-stain samples were treated with 2% uranyl acetate and subsequently dried at room temperature. Images were recorded on a FEI Tecnai Spirit microscope operating at 120 kV, equipped with a LaB₆ filament and a 2k by 2k FEI Eagle CCD camera at a magnification of 90,000.

2.6.2 Preparation of cryo-grids

Pol II-Iwr1 complexes were applied to carbon coated holey grids as described (see 2.6.1). Cryo-samples were flash-frozen in liquid ethane using a semi-automated controlled-environment system (Vitrobot, FEI Company) at 4° C, 95% humidity, and stored in liquid nitrogen until transfer to the microscope.

2.6.3 Cryo-electron microscopic data collection

Micrographs were recorded on Kodak SO-163 electron film with 100 ms pre-exposure under low-dose conditions of ~ 25 electrons/ \AA^2 on a Tecnai Polara F30 field emission gun microscope operating at 300 kV and underfocus values in the range of 1.5-3.5 μm . Negative film images were scanned on a Heidelberg drum scanner with a pixel size of 1.23 \AA on the object scale.

2.6.4 Single particle reconstruction

The contrast transfer function was determined using CTFFIND and SPIDER (Frank et al, 1996). Particles were picked automatically with SIGNATURE (Chen & Grigorieff, 2007) and subjected to manual negative selection. Data were processed using SPIDER (Frank et al, 1996). Initially 47,617 particles from 35 micrographs were decimated three-fold by pixel binning and aligned to the complete Pol II structure (1WCM) (Armache et al, 2005a) that was Gaussian filtered to 25 \AA resolution. After the first 20 rounds of angular refinement particles were sorted against two references: 1) the current reconstruction and 2) the initial reference filtered to 22.5 \AA resolution. A cross-correlation value was calculated for each particle projected onto each of the two references. Particles were assigned to the reference group bearing the higher cross-correlation value. Particle groups were then back-projected separately and the resulting reconstructions were used as input references for the next round of sorting. This

procedure was repeated until the distribution of particles converged after five more rounds of sorting and refinement. Sorting resulted in a volume based on 25,206 particles showing a significant difference density in the Pol II cleft and 22,411 particles being assigned to the initial reference. Also a significant positive difference density appeared next to the Rpb4/7 region that was complemented by a negative difference density on the opposite side of Rpb4/7. This implied a shift of this region relative to the position in the crystal structure and substantiated that the map was free of reference bias and represented experimental data. The images were back-projected in real space using the refined angles. The resulting reconstruction was Butterworth low-pass filtered to 22.5 Å and used as a reference for another round of alignment and refinement as described above. This procedure was iterated until no changes in growth of additional densities were observed. Finally, a 20.9 Å reconstruction in real space using the BP RP algorithm of SPIDER was obtained. The resolution for the final volume was estimated based on a cut-off value of 0.5 for the Fourier shell correlation. A difference density was calculated by subtracting the reference filtered to the same resolution from the Pol II-Iwr1 reconstruction using normalized maps. Additionally, a difference density with a reconstruction of Pol II by itself was calculated which was in agreement with the location of the major difference density derived by the prior analysis. Although this difference density was larger, interpretation was complicated by the presence of higher noise levels.

2.7 Gene expression profiling

2.7.1 Sample preparation and microarray measurement

For microarray measurements cells were grown in YP medium with 2% glucose. Three independent colonies were used for inoculation of overnight pre-cultures, which were diluted to an OD₆₀₀ of 0.1 and grown to 0.8 the next day (25 ml cultures, incubator shaking at 160 rpm, 30 °C). Cells were harvested by centrifugation (4,000 rpm, 3 min, 20 °C). Total RNA was prepared after cell lysis

using a mixer mill (Retsch) and subsequent purification using the RNeasy kit (Qiagen). The total RNA preparation was treated on-column with DNase (Qiagen). All following steps were conducted according to the Affymetrix GeneChip Expression Analysis Technical Manual (P/N 702232 rev. 2). Briefly, one-cycle cDNA synthesis was performed with 1 µg of total RNA. *In vitro* transcription labelling was carried out for 16 h. The fragmented samples were hybridized for 16 h on Yeast Genome 2.0 expression arrays (Affymetrix), washed and stained using a Fluidics 450 station, and scanned on an Affymetrix GeneArray scanner 3000 7G.

2.7.2 Statistical data analysis

Data analysis was performed with support from Tobias Koschubs using R/Bioconductor (Gentleman et al, 2004). *Schizosaccharomyces pombe* probes were removed before normalization with the GCRMA algorithm (Wu & Irizarry, 2004). Linear model fitting and multiple testing correction using an empirical Bayes approach was performed using the LIMMA package (Smyth, 2004). Differentially expressed genes were defined as having an adjusted P-value smaller than 0.05 and an estimated fold-change of at least 2.0 (based on the mean expression levels in the triplicate measurements). Classification of significantly differentially expressed genes was carried out according to Super GO-Slim: Process of the SGD Gene Ontology Slim Mapper (www.yeastgenome.org). To check whether the effects of an IWR1 knockout resemble those of a transcription factor knockout, Achim Tresch compared the IWR1 knockout profiles with a re-analysis (Reimand et al, 2010) of a yeast transcription factor knockout compendium (Hu et al, 2007). Pair-wise Pearson correlations of the fold expression (relative to wild type) were calculated between all profiles. The first two principal components of the resulting correlation matrix were visualized in Figure 13 (see 3.4.3).

3 Results and Discussion

3.1 RNA polymerase II nuclear localization is specifically dependent on the Iwr1 NLS

Pol II assembles from its subunits in the cytoplasm (Boulon et al, 2010). How the entire complex is imported into the nucleus is unknown. Since Pol II subunit sequences lack a known nuclear localisation signal (NLS) nuclear import apparently requires an additional factor (see 1.5.2). Analysis of the protein interactome in the yeast *Saccharomyces cerevisiae* revealed that Pol II subunits co-purify with the protein Iwr1 (Gavin et al, 2002; Krogan et al, 2006) (see 1.3). As shown in proteome-wide localization experiments, Iwr1 is localized both in the cytoplasm and in the nucleus (Huh et al, 2003), making it a likely candidate for being involved in nuclear import of Pol II.

3.1.1 Iwr1 contains a bioinformatically predicted bipartite NLS that is functional *in vivo*

Inspection and bioinformatic analysis of several Iwr1 homologues with the program cNLS Mapper (Kosugi et al, 2009b) led to the identification of a putative bipartite classical NLS (Figure 1). Earlier computational identification might have been missed due to the unusually long linker of 14 a. a. between the two basic parts of the NLS in *Saccharomyces cerevisiae* that is not accounted for by the PROSITE profile (Dingwall & Laskey, 1991).

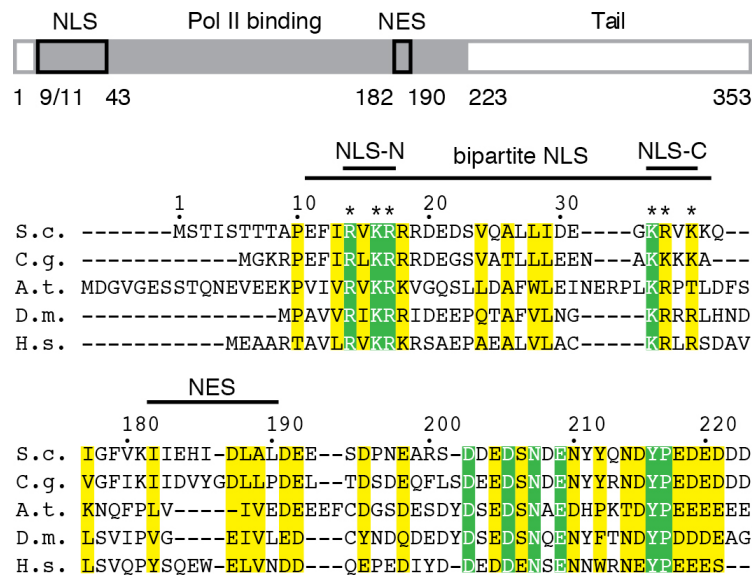


Figure 1 · Sequence conservation and functional elements of Iwr1. Schematic representation of yeast (S.c.) Iwr1 and amino acid sequence conservation. The Pol II-binding region is indicated in grey, boxes delineate NLS and NES sequences. Invariant residues are highlighted in green, residues with $\geq 80\%$ conservation in yellow. Asterisks indicate mutations to alanine in the mutant variants. S.c. *Saccharomyces cerevisiae*, C.g. *Candida glabrata*, A.t. *Arabidopsis thaliana*, D.m. *Drosophila melanogaster*, H.s. *Homo sapiens*.

While this work was ongoing, a leucine-rich nuclear export signal (NES) in Iwr1 being recognized by the karyopherin Xpo1 was reported (Peiro-Chova & Estruch, 2009). As the functional relevance of the putative Iwr1 NLS could not be shown easily by nuclear exclusion of the protein upon mutation of the NLS, the possibility was used to trap the factor within the nucleus by deleting its NES. As expected enhanced green fluorescent protein (EGFP)-tagged Iwr1 expressed from a plasmid under its natural promoter showed a pan-cellular localization, which was shifted to an exclusively nuclear localization upon deletion of its NES (Figure 2). Additionally mutating the conserved arginines and lysines to alanines in the N-terminal part of the bipartite NLS (Figure 1) abrogated nuclear trapping and consequently led to cytoplasmic localization of Iwr1 Δ NES/NLS-N_{mutant}-EGFP (Figure 2). Thus, the *in vivo* functionality of the bioinformatically predicted bipartite NLS in Iwr1 could be confirmed. Accordingly, Iwr1 is a nucleo-cytoplasmic shuttling protein by means of its own NLS and NES.

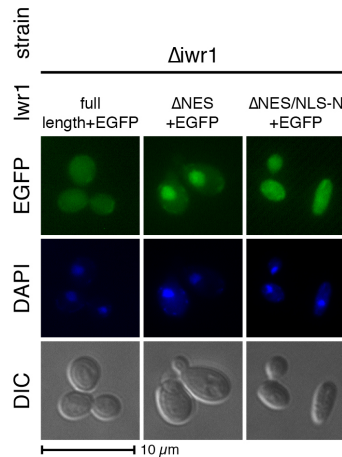


Figure 2 · Nuclear entry of Iwr1 is dependent on its bipartite NLS. Localization of Iwr1-EGFP in $\Delta iwr1$ cells expressing the indicated Iwr1 variants. Nuclear staining with DAPI.

3.1.2 RNA polymerase II mislocalizes to the cytoplasm in cells lacking Iwr1

To test the hypothesis that Iwr1 is involved in nuclear import of Pol II, yeast strains carrying an EGFP fusion of the genes encoding the Pol II subunit Rpb1, Rpb3 and Rpb4, respectively, were constructed. These subunits were chosen, as Rpb1 and Rpb3 represented two major Pol II assembly intermediates (Boulon et al, 2010) and Rpb4 was described to additionally be involved in processes independent of the core polymerase (Selitrennik et al, 2006). In growing cells from these strains, fluorescence was restricted to the nucleus, indicating that Pol II was localized in the nucleus as expected (Figure 3). However, in a strain that additionally contained a deletion of the gene encoding Iwr1 ($\Delta iwr1$), Pol II, represented by the three subunits, accumulated in the cytoplasm (Figure 3). The delocalization of Rpb4 was not as severe as for Rpb1 and Rpb3, which seems understandable as Rpb4 is described to undergo nucleo-cytoplasmic shuttling independent of the core polymerase (Selitrennik et al, 2006) and considering the up to eleven-fold higher amount of Rpb4 as compared to Rpb3 that is limiting for Pol II complex formation (Kimura et al, 2001).

In addition to the Pol II delocalization, $\Delta iwr1$ cells had an increased size and typically possessed an elongated shape. Nuclear localization of Pol II could however be rescued by transformation of a plasmid expressing Iwr1 from its natural promoter (Figure 3A). Similarly, the cells returned to their wild-type morphology.

Iwr1 is therefore required for proper nuclear localisation of Pol II.

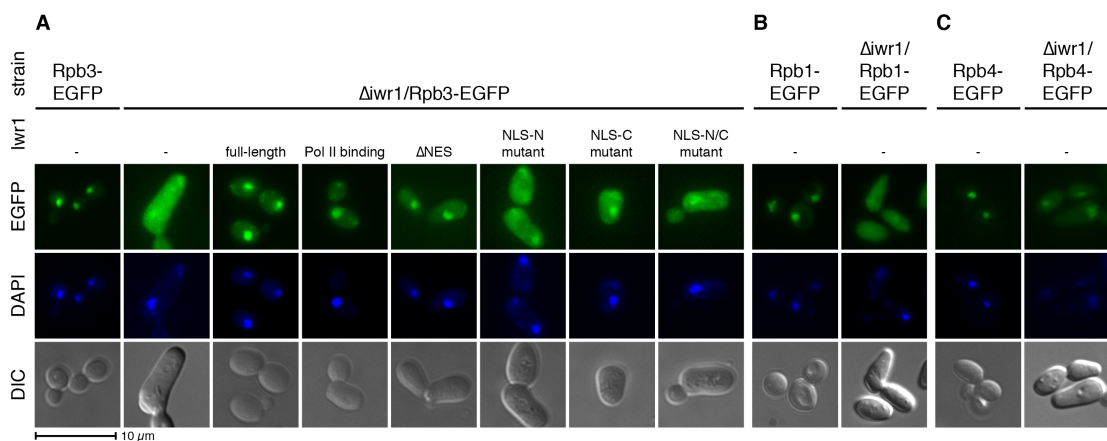


Figure 3 · Iwr1 is required for nuclear localization of Pol II. Localization of Rpb3- (A), Rpb1- (B) and Rpb4-EGFP (C) in wild-type and $\Delta iwr1$ cells expressing the indicated Iwr1 variants (-, empty plasmid). Nuclear staining with DAPI.

3.1.3 RNA polymerase II nuclear localization requires the Iwr1 NLS

Rescue of $\Delta iwr1$ cells depended on the NLS of Iwr1, as it was not achieved with Iwr1 variants that lacked the NLS (not shown) or carried mutations in NLS subsites (Figure 3A). In general, Pol II mislocalization persisted upon expression of Iwr1 variants with mutations in either the N- or C-terminal part of the bipartite NLS, as well as with both of them combined. Yet, the severity of the mislocalization seemed to vary between different NLS mutants, which is why a more detailed quantitative investigation was carried out in addition (see 3.5.1).

Surprisingly, the Iwr1 variant with the NES deletion was capable of restoring nuclear Pol II localization, although the cells seemed to have a tendency towards the more elongated shape, characteristic for the Δ iwr1 mutant.

Apparently, cells can tolerate loss of cyclic shuttling of Iwr1 between the cytoplasm and the nucleus during exponential growth. In this view, nuclear import of one Pol II complex per Iwr1 protein is sufficient to bring about a close-to-wild-type phenotype. However, the function of the Iwr1 NLS is indispensable for efficient Pol II nuclear import.

3.1.4 The human homologue of Iwr1 partially rescues RNA polymerase II nuclear localization in *Saccharomyces cerevisiae*

To assess the evolutionary relevance of Iwr1-dependent Pol II nuclear import beyond yeast, an in-depth search for homologues was undertaken. Using the sensitive CSI-BLAST method (Biegert & Soding, 2009) a putative homologue in humans could be found (Figure 1). Most importantly, the NLS sequence is conserved from yeast to humans. No homologues were detected in bacteria or archaea.

The functional conservation of the human Iwr1 homologue was assessed using the established Pol II localization assay in Δ iwr1 yeast cells (Figure 4). A synthetic gene encoding human Iwr1 with codons optimized for expression in *Saccharomyces cerevisiae* (see 2.2.2) was inserted under the control of the yeast Iwr1 promoter and terminator sequences in a rescue plasmid. After transformation of the plasmid into Δ iwr1 yeast cells Pol II nuclear localization was indeed partly restored.

Therefore, human Iwr1 is functional in yeast, although the conservation of 39% (conserved residues) between the yeast and corresponding human Pol II binding region (see 3.2.2) is rather low. More importantly, the mechanism elucidated for yeast Pol II nuclear import in this work, is likely to hold true in humans.

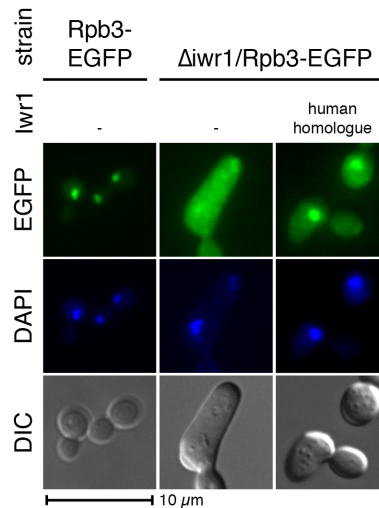


Figure 4 · The human homologue of yeast *Iwr1* partially rescues the growth defect in $\Delta iwr1$ cells. Localization of Rpb3-EGFP in wild-type and $\Delta iwr1$ cells expressing the human *Iwr1* homologue (-, empty plasmid). Nuclear staining with DAPI.

3.1.5 RNA polymerase II nuclear localization is independent of Rpb4

Although no NLS was found in any of the Pol II core subunits, a Pol II independent role for the Rpb4/7 heterodimer involving nucleo-cytoplasmic shuttling was reported as already mentioned (Selitrennik et al, 2006). Yet, no NLS was described in the Rpb4/7 heterodimer by this study.

A closer bioinformatic inspection of the Rpb4 sequence revealed a potential bipartite N-terminal NLS with a good confidence score ranging from a. a. 8 to 40 (Kosugi et al, 2009b). As RPB4 is not essential the gene was deleted in the Rpb3-EGFP reporter strain. Pol II localization was unaltered and remained exclusively nuclear (Figure 5). The parallel deletion of IWR1 and RPB4 resulted in the typical $\Delta iwr1$ phenotype, both in terms of Pol II delocalization and growth defect, also confirming earlier indications of a suppression of the RPB4 phenotype upon deletion of IWR1 that remains unexplained (Collins et al, 2007).

A role of the nucleo-cytoplasmic shuttling of the Rpb4/7 heterodimer in nuclear import of Pol II can thus be discarded further strengthening the case for *Iwr1* being the principal factor for Pol II nuclear import.

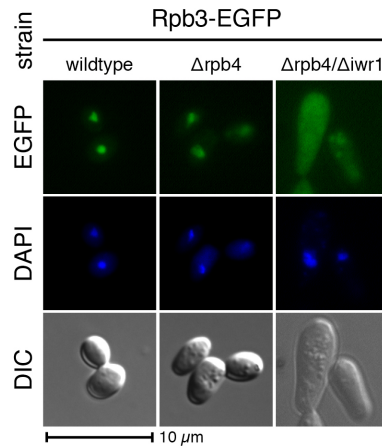


Figure 5 · Pol II nuclear import is independent of Rpb4. Localization of Rpb3-EGFP in wild-type, $\Delta rpb4$ and $\Delta rpb4/\Delta iwr1$ cells. Nuclear staining with DAPI.

3.1.6 RNA polymerase I and III nuclear localization are independent of Iwr1

To assess the role of Iwr1 in nuclear import of Pol I and Pol III, the two largest subunits of these polymerases, Rpa190 and C160, respectively, were genomically tagged with EGFP. From these strains, variants were derived that additionally contained a genomic deletion of IWR1.

As expected, Rpa190 shows a nucleolar¹ and C160 a nuclear localization in wild-type background (Figure 6). In the strains with an IWR1 deletion the localization of these two subunits was not altered. Although there seemed to be a slight increase in fluorescent signal for C160 in the cytoplasm, the change was by far not as severe as for Rpb1 or Rpb3 representing Pol II. As association of Iwr1 to Pol I or Pol III could neither be confirmed *in vivo*² (Gavin et al, 2002; Krogan et al, 2006) nor *in vitro* (see 3.2.1) the slight delocalization might be due to an indirect effect. This seems plausible when the broad changes in gene expression

¹ The nucleolar localisation is not well resolved in Figure 6, although it was clearly visible by eye under the microscope.

² Krogan et al, 2006 found a low-confidence interaction with AC40, a subunit that is shared between Pol I and Pol III (see Table 1).

in the $\Delta iwr1$ strain are considered that might lead to indirect alterations in the cell.

Based on these findings the conclusion is drawn that Iwr1 is not involved in nuclear import of the “odd” polymerases Pol I and Pol III and hence that Iwr1 is a specific factor for Pol II nuclear import.

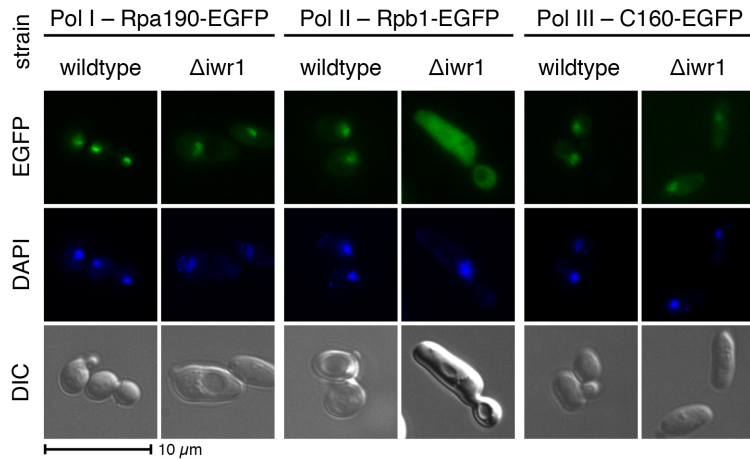


Figure 6 · Pol I and III nucleolar and nuclear localization, respectively are independent of Iwr1. The localization of Pol I, II and III largest subunits Rpa190-EGFP, Rpb1-EGFP and C160-EGFP, respectively, in cells with and without IWR1 is depicted. Nuclear staining with DAPI.

3.2 Iwr1 binds RNA polymerase II directly via an N-terminal binding region

Co-purification of Iwr1 with Pol II subunits does not necessarily imply a direct interaction. Additional Pol II-related proteins, such as Tfg1, Tfg2 and Spt5, as well as other proteins, such as Cbf2 (a kinetochore protein) or Cns1 (a HSP90 co-chaperone) were identified in the complexes (Krogan et al, 2006), which might provide a physical link between Iwr1 and Pol II. To clarify this ambiguity the formation of the Pol II-Iwr1 complex was recapitulated *in vitro*.

3.2.1 Iwr1 binds directly and specifically to RNA polymerase II

Recombinantly expressed and purified yeast Iwr1 was incubated in five-fold excess with pure Pol II. Then, the sample was passed over a size exclusion chromatography column and the peaks and fractions were analyzed by SDS-PAGE (Figure 7A) and static light scattering measurements. Iwr1 co-eluted with the other Pol II subunits in a high-molecular weight peak, indicative of stable binding of the factor. The identity of the Iwr1 band was confirmed by mass spectrometry. Moreover, the molecular weight of the protein complex based on the static light scattering analysis was around 540 kDa, as compared to 514 kDa for Pol II only, which was in support of complex formation.

Also, in the same manner, binding to Pol I and Pol III was assessed. As only degradation products of larger polymerase subunits could be detected by mass spectrometric analysis of SDS-PAGE bands in the region where Iwr1 was to be expected, even a weak, sub-stoichiometric association seems unlikely (not shown).

Iwr1 therefore binds directly and specifically to Pol II. No post-translational modifications of the protein, such as a possible phosphorylation at serine 313 (Albuquerque et al, 2008), are relevant for the association.

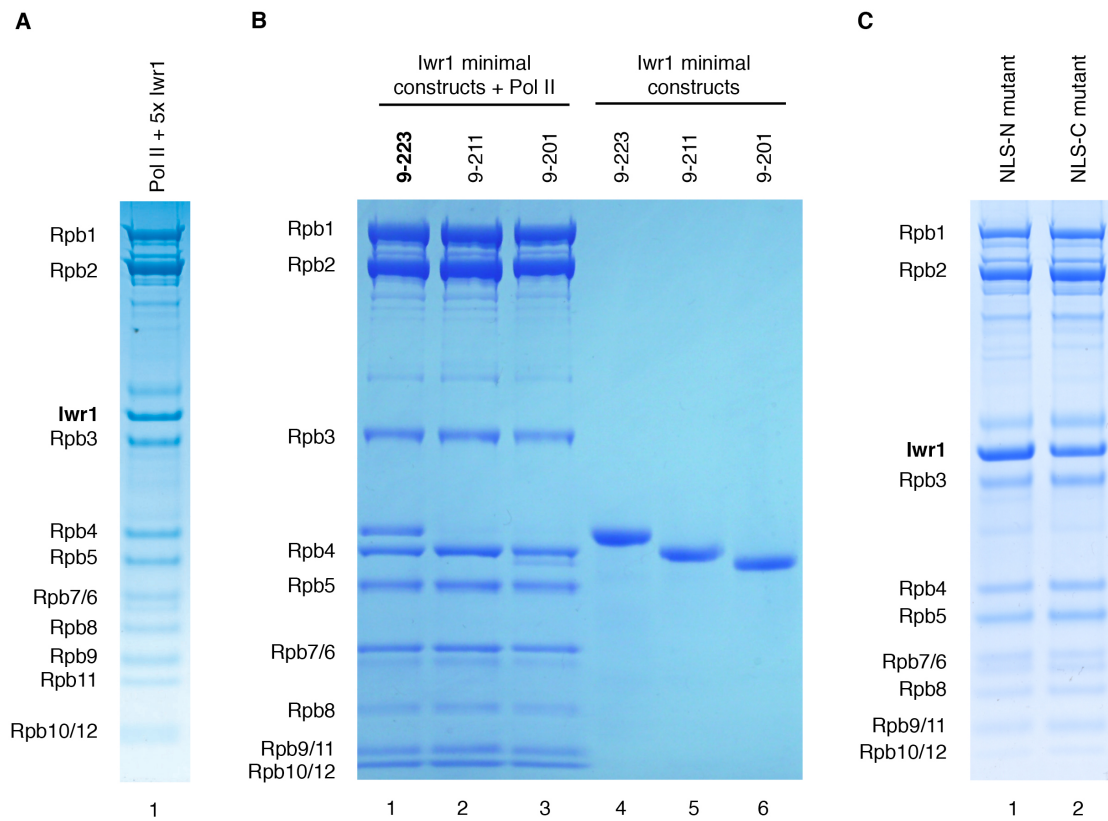


Figure 7 · Assemblies of Iwr1 variants with Pol II. Coomassie-stained SDS-PAGEs after size-exclusion chromatography (A-C). (A) Assembly of full-length Iwr1 with Pol II. (B) Determination of minimal Pol II binding constructs. Assemblies with Pol II [lanes 1-3] and purified recombinant Iwr1 variants (lanes 4-6). Iwr1 (9-223) binds stoichiometrically (defined as the minimal Pol II binding region); Iwr1 (9-211) binds, yet stoichiometry cannot be determined due to an overlap of the respective band with Rpb4; Iwr1 (9-201) binds substoichiometrically. (C) Assembly of Pol II with Iwr1 NLS-mutants N and C.

3.2.2 Iwr1 contains a minimal RNA polymerase II-binding region that is sufficient for nuclear localization of the enzyme

Using N- and C-terminally truncated variants of full-length Iwr1 in Pol II binding assays a minimal Pol II binding region of the protein was delineated (Table 12, Figure 7B). Truncation at the C-terminus until position 223 did not affect Iwr1 expression levels and stoichiometric binding to Pol II. Further C-terminal truncation lead to an overlap of the bands of Iwr1 [9-211] and Rpb4. Although this band seemed to be stronger than expected for a single protein, full stoichiometric binding could not be confirmed unambiguously. C-terminal

truncation to position 201, decreased expression levels and severely affected Pol II binding. At the N-terminal end a truncation to position 9 was possible without any detrimental effect on expression and binding, whereas further shortening to position 19 almost abolished binding to Pol II. However, a construct starting only at position 55 was still capable of showing a similarly weak interaction with Pol II.

A stable and well-expressed minimal Pol II-binding region was therefore defined as extending from position 9 to 223. Interestingly, the boundaries of this construct coincided with and included the bipartite NLS at the N-terminus and another strongly conserved region at the C-terminus (Figure 1).

In *in vivo* rescue experiments of the Δ *lwr1* strain, the minimal Pol II-binding region turned out to be capable of an almost wild-type-like restoration of Pol II nuclear localization (Figure 3A). The lowly conserved C-terminal third of *lwr1* therefore seems dispensable in terms of both binding to and nuclear import of Pol II.

Table 12 · Constructs used for delineation of the *lwr1* minimal Pol II-binding region

Iwr1	N-terminus (extending to a. a. 353)			C-terminus (starting at a. a. 1)		
	9	19	55	201	211	223
position [a. a.]	9	19	55	201	211	223
expression	good	bad	bad	bad	good	good
Pol II binding	yes	weak	weak	weak	possibly	good

3.2.3 RNA polymerase II-binding is not affected by mutations of the *lwr1* NLS

The *lwr1* NLS mutants are not capable of restoring Pol II nuclear localization in the Δ *lwr1* strain (see 3.1.3, Figure 3A). The most likely interpretation of this finding is to infer impaired NLS functionality. Alternatively, these mutations could impair Pol II binding and as a consequence result in non-recruitment of required factors for Pol II assembly or import. To exclude this possibility the *lwr1* NLS N and NLS C mutants were recombinantly expressed and purified in order to assess their capability of Pol II-*lwr1* complex formation.

As expected, both mutants behaved like the wild-type protein during the purification. No aggregate peak at the exclusion volume was detectable, indicating proper protein folding. Furthermore, degradation was present in similarly low amounts as for wild-type Iwr1. Most importantly, both NLS mutants formed stoichiometric complexes with Pol II (Figure 7C).

Iwr1 NLS mutants therefore should still be functional in recruiting hypothesized other factors of relevance to Pol II. Only binding of karyopherins that directly recognize the NLS should be affected by the point mutations.

3.3 Iwr1 binds in the RNA polymerase II active centre cleft

Knowing about the architecture of protein complexes is a valuable starting point to infer possible mechanisms for a certain factor. As an atomic model was available for Pol II it was desirable and for a first model sufficient to locate Iwr1 on Pol II. Therefore, the Pol II-Iwr1 complex was analysed by cryo-electron microscopy (EM) and single particle reconstruction.

3.3.1 A major cryo-EM difference density locates Iwr1 in the RNA polymerase II active centre cleft

To assess the homogeneity of the Pol II-Iwr1 complex samples were negatively stained with Uranyl acetate. Single roundish particles that well represented polymerases were visible on the grid (Figure 8B). Next, complexes were embedded in vitreous ice and images were recorded under low electron dose and cryo-conditions (see 2.6). Starting with 47,617 particles and sorting against the initial reference after the first 20 rounds of angular refinement a final volume based on 25,206 particles at 21 Å was obtained. The resolution was estimated based on a conservative cut-off value of 0.5 for the Fourier shell

correlation. Calculating a difference density map between the Pol II-Iwr1 reconstruction and the complete Pol II crystal structure (Armache et al, 2005b) resulted in two significant positive densities (Figure 8A). One of them appeared in the Pol II cleft, which is formed by the two largest subunits Rpb1 and Rpb2. The other one was on the upstream side of the Rpb4/7 heterodimer, complemented by a negative difference density on the downstream side. This implied a shift of this region relative to the position in the crystal structure and substantiated that the map was free of reference bias and represented experimental data. This flexibility can be rationalized as the Rpb4/7 heterodimer contacts the polymerase core via the Rpb1 clamp region that has been shown to be flexible in the first Pol II crystal structure (Cramer et al, 2000). Also underpinning this flexibility, the density at the Rpb4/7 heterodimer was smaller in cryo-EM reconstructions of Pol II alone than in reconstructions of Pol II with nucleic acids that bind in the cleft and thus lock the clamp in a certain position. It was therefore assumed that the difference density within the cleft was due to the presence of Iwr1 in the complex. As the size of the density does not seem to be sufficient to account for the full molecular weight of protein, parts of Iwr1 need to be flexible. This is not unlikely as the C-terminal third of the protein can be truncated without affecting the stability of the Pol II-Iwr1 complex in size exclusion experiments and therefore is probably not contacting the polymerase. Moreover, parts of the N-terminal NLS might also be flexible, as they need to be accessible to karyopherins.

Overall, the cryo-EM reconstruction of the Pol II-Iwr1 complex indicated that Iwr1 binds Pol II between its two largest subunits Rpb1 and Rpb2 within the active centre cleft.

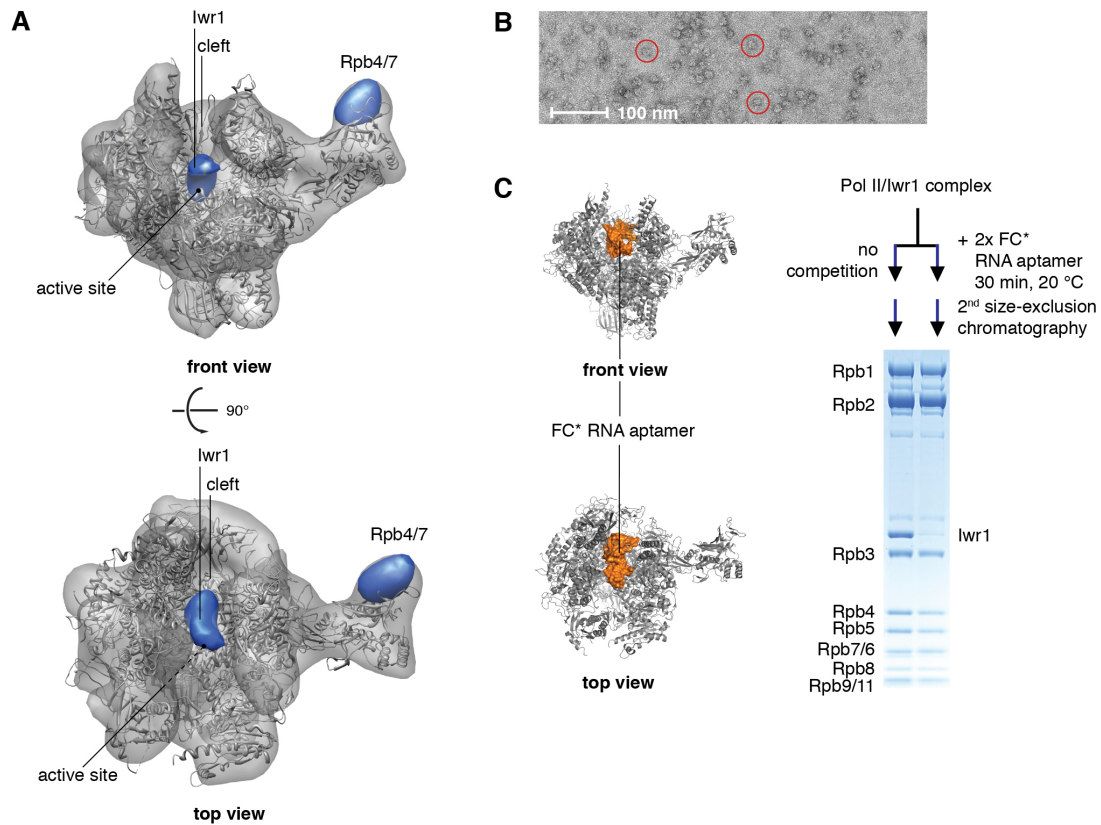


Figure 8 · Iwr1 binds Pol II in the active centre cleft. (A) Cryo-EM map of the Pol II-Iwr1 complex at 21 Å resolution (grey). The difference density between this map and a Pol II crystal structure (PDB 1WCM), both filtered to 25 Å, is in blue. A difference density calculated with a free Pol II cryo-EM reconstruction is consistent, albeit noisier (not shown). Pol II views and regions as previously defined (Cramer et al, 2001). Figures prepared with UCSF Chimera. (B) Electron micrograph of Pol II-Iwr1 complexes negatively stained with uranyl acetate. Red circles indicate distinct Pol II-Iwr1 complexes. (C, left) FC* as located in the X-ray structure is in orange (PDB 2B63). Figures prepared with PyMOL (Schrödinger, LLC). (C, right) Coomassie-stained SDS-PAGE of the Pol II-Iwr1 complex after size-exclusion chromatography, before (left), and after (right) competition with the FC* RNA aptamer (Kettenberger et al, 2006). The identity of the bands was confirmed by mass spectrometry.

3.3.2 Competition experiments with nucleic acids confirm the predicted location of Iwr1 on RNA polymerase II

To confirm the architecture of the Pol II-Iwr1 complex inferred from cryo-EM and single particle analysis, competition experiments with nucleic acids were carried out. As the precise location of a minimal scaffold is known from structural studies on the Pol II elongation complex (Kettenberger et al, 2004) it is suited to probe for binding of a factor in this region. Similarly, a structure of an

RNA aptamer bound to Pol II is available (Kettenberger et al, 2006). As the aptamer is small (10.7 kDa) its spatial location is confined to only a small region directly above the bridge helix. Importantly, the structure of polymerase itself remains unaffected by the presence of either a minimal scaffold or the RNA aptamer, so that an allosteric mechanism of displacement can be excluded.

Adding either the RNA aptamer (Figure 8C) or a minimal scaffold (Figure 10E) together with Iwr1 to Pol II resulted in formation of the respective Pol II nucleic acid complexes. In size exclusion experiments the 260 to 280 nm absorption ratio in the Pol II peak was increased as compared to the ratio in absence of nucleic acids. Also in a SDS-PAGE of these fractions the band corresponding to Iwr1 almost completely disappeared (Figure 8C and Figure 10C). Introducing a specific order of addition of Iwr1 and nucleic acids did not affect the final complex obtained, indicating that the chemical binding equilibrium between Iwr1, nucleic acids and the Pol II complex is reached fast as compared to the 30 min allowed for complex assembly (Figure 10C).

As a result, both the RNA aptamer and a DNA/RNA scaffold are capable of displacing Iwr1 from Pol II confirming – with their defined spatial location – the architecture of the Pol II-Iwr1 complex derived by cryo-EM. As a functional consequence *in vivo*, the presence of Iwr1 is incompatible with a DNA template strand loaded into the Pol II active centre cleft. Presence of the factor on chromatin is therefore not to be expected.

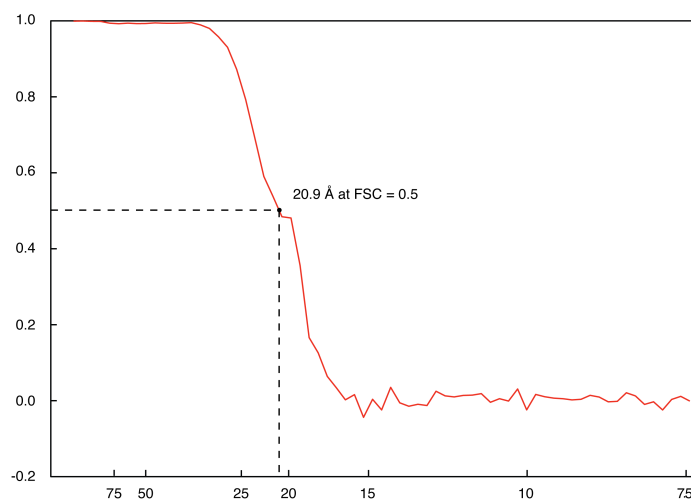


Figure 9 · Fourier shell correlation (FSC) curve of the Pol II-Iwr1 cryo-EM map. The resolution at 0.5 is indicated.

closed DNA-TBP-TFIIB complex as indicated. The binding of DNA to Pol II was detected by the increased absorption at 260 nm in the Pol II size-exclusion chromatography peak. (B) Competition of Iwr1 and TFIIB for binding to Pol II. Iwr1 (5-fold excess) was incubated with Pol II for 20 min (lane 1). TFIIB (5-fold excess) was additionally added for 20 min (lane 2). (C) Competition of Iwr1 and a minimal DNA/RNA nucleic acid scaffold for binding to Pol II. Either Iwr1 (5-fold excess) was pre-incubated with Pol II for 20 min and then the DNA/RNA scaffold (2-fold excess) was added for 20 min (lane 1) or vice versa (lane 2). The binding of DNA/RNA to Pol II was detected by the increased absorption at 260 nm in the Pol II size-exclusion chromatography peak. (D) Sequence of the closed promoter DNA used in (A). BREu/d, upstream/downstream TFIIB recognition element. (E) The minimal DNA/RNA scaffold used in (C). The bridge helix indicates the position at which the scaffold is located when bound to Pol II (Damsma et al, 2007).

3.3.4 Iwr1 hampers the crystallization of RNA polymerase II

Beyond an architectural model of the Pol II-Iwr1 complex an atomic structure would be desirable. To this end crystallization of the complex in established Pol II crystallization conditions was tried (Armache et al, 2005b; Kettenberger et al, 2006). Crystal growth of Pol II-Iwr1 was delayed as compared to formation of crystals from Pol II alone and the final crystals in the Pol II-Iwr1 setup were small with most of them below 100 μm in diameter. Additionally, precipitate was present in the crystallization drops that was likely due to Iwr1 protein (Figure 11). A diffraction pattern for one of the small crystals could however be collected and an electron density map calculated. Yet, no density in addition to Pol II was present. Possibly the altered position of the Rpb4/7 heterodimer seen in the cryo-EM reconstruction (Figure 8A) or parts of Iwr1 would lead to clashes within the Pol II crystal lattice that is formed in the crystallization conditions used. A *de novo* crystallization screening led to the identification of novel conditions, which were akin to conditions already known. Macroscopically, the crystals in the screen resembled typical Pol II crystals so that optimization was not pursued into this direction.

After these trials, the most promising approach to crystallize the Pol II-Iwr1 complex seems to be to use a minimal functionally relevant Iwr1 truncation variant or the Pol II core instead of the complete polymerase (see 4.2).

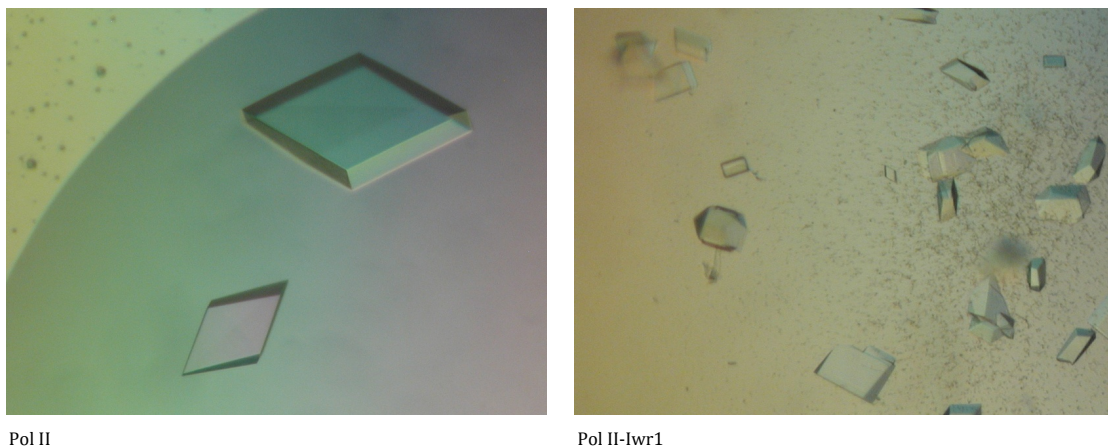


Figure 11 · Crystallization of Pol II and trials with Pol II-Iwr1. Protein was concentrated to around 4 mg/ml and hanging drop crystallization was set up in 200 mM ammonium acetate, 300 mM sodium acetate, 50 mM Hepes, pH 7.0, 5% PEG 6,000, 5 mM TCEP. Crystal growth in the case of Pol II-Iwr1 was delayed for about three days. In addition, precipitate appeared.

3.4 Iwr1 does not seem to be involved in transcription

While this thesis was ongoing, a study was published that tried to establish Iwr1 as a transcription factor (Peiro-Chova & Estruch, 2009). Knowing about the relevance of Iwr1 for Pol II nuclear import, its direct interaction with Pol II that would seem unusual for a transcription factor, the architecture of the complex and the competition with nucleic acids, a further characterization of Iwr1 regarding a direct involvement in transcription was required.

3.4.1 Iwr1 is not required for promoter-dependent transcription *in vitro*

To assess the effect of Iwr1 on transcriptional initiation an *in vitro* transcription system based on yeast nuclear extracts, recombinant Gal4-VP16 transcriptional activator and a plasmid with the yeast HIS4 core promoter containing an upstream Gal4 binding site was used. Nuclear extracts were prepared from wild-type and $\Delta iwr1$ cells. Both supported transcriptional initiation, with the $\Delta iwr1$ extracts at possibly slightly higher levels as the

wild-type extract (Figure 12). Addition of recombinant Iwr1 protein led to a significant decrease in transcriptional signal, whereas addition of control protein did not.

Therefore, Iwr1 is generally not required for transcriptional initiation but rather interferes with it. Possibly, binding of Iwr1 to Pol II reduces the concentration of free polymerases available for initiation and thereby exerts its inhibitory effect.

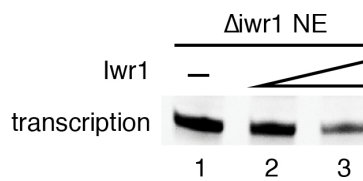


Figure 12 · Nuclear extracts from $\Delta iwr1$ cells support transcriptional initiation. Bands indicate the transcriptional signal. In lanes 2 and 3, 1 and 10 pmol, respectively, of recombinant Iwr1 were added to the reaction.

3.4.2 Iwr1 is not present at chromatin

According to the competition of Iwr1 with nucleic acids in binding to Pol II and considering the architecture of the Pol II-Iwr1 complex the protein would not be expected to be bound in the context of chromatin. Nevertheless, a different mode of binding or a state in which Pol II is already partly recruited to a promoter but has not loaded the DNA template strand into its active centre might allow for the presence of Iwr1 at chromatin. To assess these possibilities we tried to co-immunoprecipitate (ChIP) genomic DNA with a C-terminally TAP-tagged version of Iwr1. After the immunoprecipitation with IgG beads DNA levels were comparable to the wild-type control. Probing for the presence of Iwr1 at the 5', ORF- and 3'-region of the ADH1, ACT1 and PMA1 genes by real time PCR showed no enrichment of these DNA regions in the sample. This finding is in agreement with others that additionally varied the affinity tag and tested both N- and C-terminal fusions (Peiro-Chova & Estruch, 2009). Moreover, the C-terminal

TAP-tag should be accessible as it was used in the original identification of Iwr1 as a Pol II-interacting factor (Krogan et al, 2006). However, this finding cannot be easily reconciled with the reported co-localization of Pol II and the *Drosophila melanogaster* homologue of Iwr1 on polytene chromosomes (Krogan et al, 2006).

In conclusion, Iwr1 is not present at chromatin, which is in line with its inhibitory effect on transcriptional initiation (see 3.4.1), its competition with nucleic acids (see 3.3.2) for Pol II binding and the architecture of the Pol II-Iwr1 complex (see 3.3.1).

3.4.3 The gene expression profile of the Δ iwr1 strain is dissimilar as compared to profiles of 263 transcription factor gene deletion strains

To infer a possibly specific role of Iwr1 in certain biochemical processes or pathways the gene expression profile of the Δ iwr1 strain was recorded. Compared to wild-type cells 794 genes had significantly changed expression levels (fold-change >2, p-value <0.05) with 40% (316 genes) being up- and 60% (478 genes) being downregulated. The 20 most strongly up- and downregulated genes are shown in Table 13. Generally, the higher fraction of downregulated genes suggests a rather activatory role for Iwr1. Clustering the affected genes according to gene ontology annotations reveals that many metabolic processes such as amino acid synthesis and energy related processes (ATP synthesis, NAD metabolism) are more strongly affected with a tendency towards downregulation of the genes involved. When looking at general biological processes in which the affected genes are involved, the cellular effects again appear to be very broad (Table 14). In conclusion, no specific involvement in a single process could be identified.

Table 13 · Significantly differentially expressed genes in $\Delta iwr1$ cells. Absolute \log_2 fold-change ≥ 1 , p-value < 0.05 . The 20 most strongly up- and downregulated genes out of 794 are shown representatively.

Upregulated (316 total, 40%)			Downregulated (478 total, 60%)		
ORF	Gene name	\log_2 fold-change	ORF	Gene name	\log_2 fold-change
YLR343W	GAS2	6,96	YGL089C	MF(ALPHA)2	-9,88
YKR046C	PET10	6,39	YER011W	TIR1	-9,77
YDR216W	ADR1	6,01	YAR071W	PHO11	-9,67
YOL058W	ARG1	5,89	YDR534C	FIT1	-8,49
YKL109W	HAP4	5,88	YBR093C	PHO5	-8,43
YFL014W	HSP12	4,83	YHR216W	IMD2	-8,36
YKR039W	GAP1	4,81	YHR136C	SPL2	-8,12
YJL088W	ARG3	4,70	YCL064C	CHA1	-8,02
YEL011W	GLC3	4,39	YDR033W	MRH1	-7,58
YKL161C	YKL161C	4,32	YAR068W	YAR068W	-7,52
YLR142W	PUT1	4,24	YAR073W	IMD1	-7,02
YER081W	SER3	3,94	YLR231C	BNA5	-6,86
YJL102W	MEF2	3,93	YJR025C	BNA1	-6,82
YPL152W	RRD2	3,84	YDL115C	IWR1	-6,13
YKR091W	SRL3	3,82	YHL047C	ARN2	-6,11
YDR042C	YDR042C	3,76	YCL027W	FUS1	-6,07
YPR002W	PDH1	3,73	YJL157C	FAR1	-6,02
YOR065W	CYT1	3,51	YIL011W	TIR3	-6,01
YDL181W	INH1	3,45	YLR136C	TIS11	-5,77
YBR230W-A	YBR230W-A	3,42	YOR009W	TIR4	-5,76

Table 14 · Classification of significantly differentially expressed genes to general biological processes according to the Super GO-Slim: Process of the SGD Gene Ontology Slim Mapper.

GOID	GO term	Frequency (out of 794)		Genome Frequency (out of 6310)	
		[total]	[%]	[total]	[%]
9987	cellular process	549	69.1	4835	76.6
8152	metabolic process	412	51.9	3595	57.0
8150	biological process unknown	198	24.9	1247	19.8
6810	transport	127	16.0	1034	16.4
6350	transcription	87	11.0	610	9.7
7049	cell cycle	57	7.2	521	8.3
6520	cellular amino acid metabolic process	41	5.2	207	3.3
23046	signaling process	33	4.2	239	3.8
-	other	7	0.9		
-	not yet annotated	4	0.5		

To further rationalize the data pairwise Pearson correlations of the fold-expressions were calculated between the $\Delta iwr1$ profile and a set of 263 profiles from transcription factor deletion strains (Hu et al, 2007; Reimand et al, 2010). In a visualization of the first two principal components of the correlation matrix the $\Delta iwr1$ profile stands out as an outlier (Figure 13). As proximity in this plot corresponds to similarity in transcriptional profiles the role of *Iwr1* can be interpreted as being distinct to all transcription factors considered.

According to its gene expression profile *Iwr1* thus has a pleiotropic role rather than functioning as a specific transcription factor. This finding is in agreement with the proposed role for *Iwr1* to be a principal factor in Pol II nuclear import.

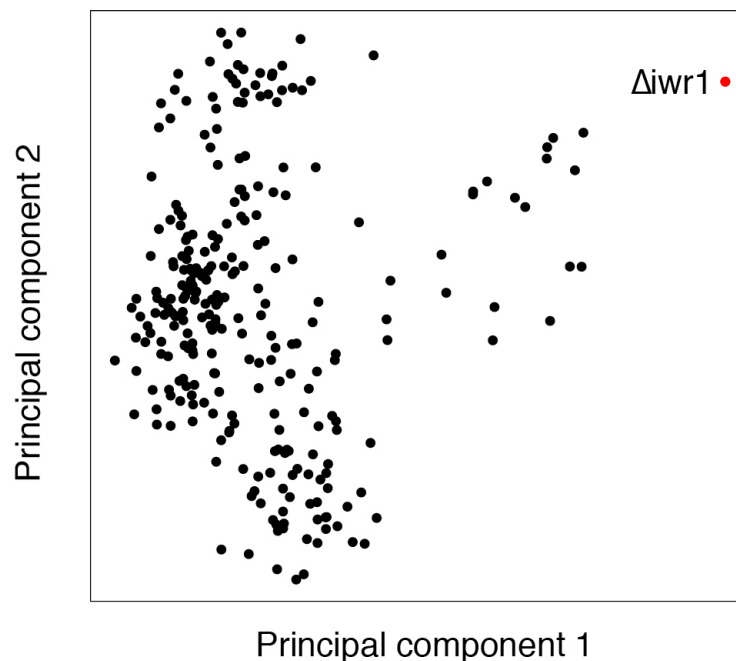


Figure 13 · The $\Delta iwr1$ gene expression profile is distinct as compared to 263 profiles of transcription factor deletions. The correlation between the profiles (relative to wild-type) was calculated. The first two principal components of the correlation matrix are plotted. $\Delta iwr1$ is indicated as a red dot.

3.5 Nuclear import of RNA polymerase II is primarily dependent on Iwr1 and yet redundant

3.5.1 Cellular growth rate is independent of RNA polymerase II nuclear localization above a critical threshold

Different Iwr1 variants led to varying degrees of Pol II nuclear localization when expressed from a rescue plasmid. To assess the relevance of proper Pol II nuclear localization for the growth fitness of rescued cells, both their nuclear to cytoplasmic fluorescence ratio of Rpb3-EGFP and their doubling time was quantified (Figure 14). As expected the wild-type Iwr1 rescue had the highest nuclear Pol II localization and the shortest doubling time whereas the control cells without a rescue had much less nuclear Pol II relative to the cytoplasmic amount and grew slower. Interestingly, the Iwr1 minimal Pol II binding region was almost as competent as the full-length protein both in restoration of growth fitness and Pol II nuclear localization. Although the Iwr1 Δ NES and NLS-C mutant were significantly worse in mediating Pol II nuclear import their doubling time was at wild-type levels. Apparently, the cells can tolerate a certain level of impaired polymerase import without severe consequences for their fitness. The critical threshold seems to be between the import efficiency of the deletion strain and the NLS-C mutation. Moreover, expression of the Iwr1 NLS-N and NLS-N/C mutant led to Pol II nuclear localization levels and growth fitness below Δ iwr1 control cells without rescue. This finding might be explained by import-incompetent Iwr1 binding to Pol II and interfering with redundant import pathways (see 4.2). As Iwr1 seems to affect this alternative pathway, it is likely to be mediated by a factor that competes with Iwr1 for Pol II binding. A candidate for this would be TFIIB for which an NLS and the responsible karyopherin Kap114 have been described (Hodges et al, 2005).

The sharp increase in doubling time between the Iwr1 NLS-C mutant rescue and the Δ iwr1 control cells lacking Iwr1 completely might also imply a second function of the protein before mediating import, possibly in polymerase biogenesis. This second function could be masked by the severe impairment of Pol II import by Iwr1 NLS-N and NLS-N/C mutants. Although these NLS mutants

should be capable of carrying out this hypothesized function the fitness of the cells ends up worse than that of $\Delta iwr1$ cells as polymerase import is below a critical threshold.

Pol II nuclear import is primarily dependent on Iwr1 as specific mutations in its localization signals result in altered Pol II localization. The presence of at least one alternative import pathway that might be mutually exclusive with Iwr1 is clear given the viability of $\Delta iwr1$ cells. However, alternative pathways might play insignificant roles in the presence of Iwr1 and only become relevant in its absence. Any sufficiently strong Pol II-interacting factor with an NLS of its own or as part of an actively imported complex might be a candidate to mediate fractional Pol II nuclear import (see 4.2).

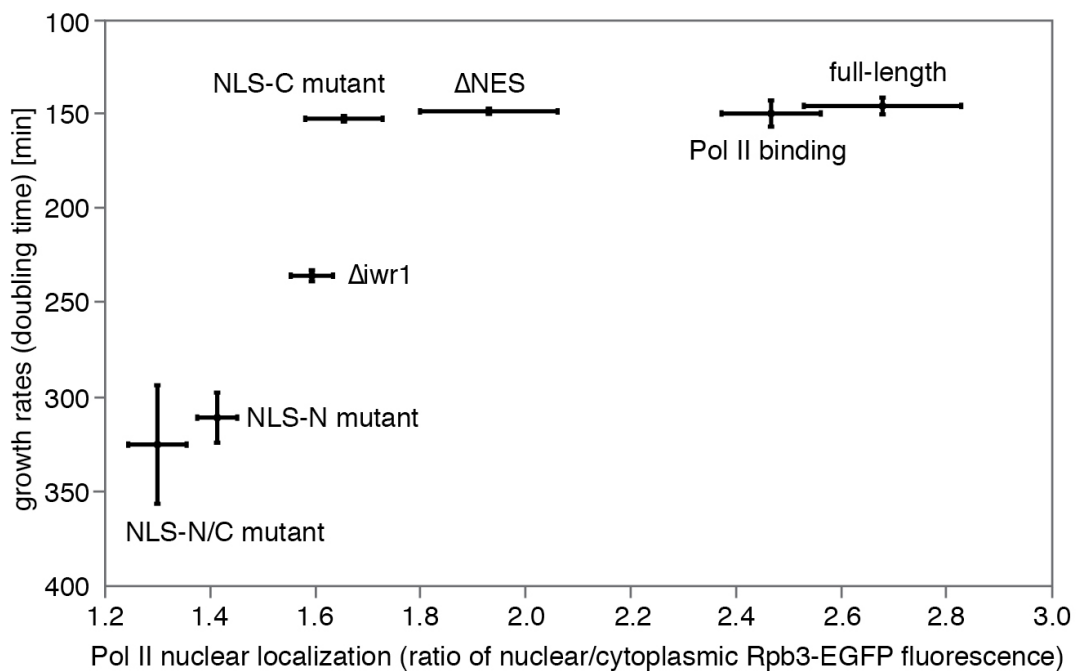


Figure 14 · Growth fitness and degree of nuclear Pol II localization show a non-linear relation. Data from $\Delta iwr1$ /Rpb3-EGFP cells expressing the indicated Iwr1 variants. Mean ratio of quantified nuclear/cytoplasmic Rpb3-EGFP fluorescence ($n \geq 10$ cells, \pm SEM) plotted against growth rates (triplicates from three independent cultures, \pm SEM).

3.5.2 Rpb1 levels and protein half-life are increased in $\Delta iwr1$ cells

Altered localization of Pol II might lead to higher total Pol II protein levels. Firstly, these would be favoured by the need to overcome less efficient nuclear import. Secondly, in wild-type cells degradation of Pol II is expected to take place in the nucleus (see 1.4.2). Therefore, cytoplasmic Pol II might be protected from targeted degradation, resulting in prolonged protein half-lives.

To test this hypothesis, total Rpb1 levels were measured in wild-type and $\Delta iwr1$ cells by quantitative western blots (Figure 15). Indeed, Rpb1 levels in $\Delta iwr1$ cells were about 8-fold higher as compared to wild-type (Figure 15A). As none of the Pol II subunit genes was significantly more highly expressed according to the $\Delta iwr1$ gene expression profile this was not the result of a specific transcriptional response of the cells to insufficient nuclear Pol II levels. To assess if Rpb1 degradation was reduced the protein half-life in a cycloheximide chase experiment was measured (Figure 15B and C). In wild-type cells the calculated Rpb1 half-life was around 25 min, which was in agreement with published results (Belle et al, 2006). However, in $\Delta iwr1$ cells no significant Rpb1 degradation could be observed over the 45 min course of the measurement.

Increased Rpb1, and possibly Pol II levels, in cells lacking *Iwr1* are due to an increased protein half-life and not due to increased expression. Considering that a large Pol II fraction is localized in the cytoplasm this finding is consistent with Pol II assembly in the cytoplasm and degradation in the nucleus.

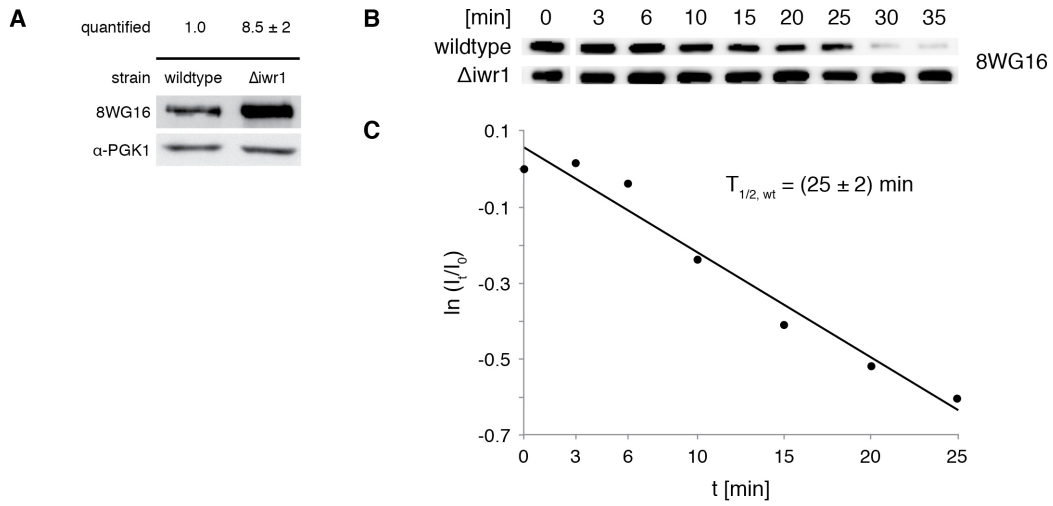


Figure 15 · Rpb1 levels and protein half-life are increased in $\Delta iwr1$ cells. (A) Quantitative western blots for Rpb1 (8WG16) and Pgk1 (α -Pgk1). Rpb1 intensity values are normalized to Pgk1 and relative to wild-type cells (triplicates, \pm SEM). (B, C) Protein half-life determination of Rpb1 in the wild-type and $\Delta iwr1$ strain. (B) Quantitative western blots for Rpb1 (8WG16). Time points after addition of cycloheximide are indicated. (C) Quantification of wild-type band intensity in (B) and Rpb1 half-life. Intensities (I) are normalized to the intensity at time point zero (I_0). Time points outside of the linear range of the antibody are not considered. The standard error of the linear regression is indicated.

4 Conclusions and Outlook

4.1 A model for RNA polymerase II nuclear import

The results of this thesis and published data converge on a cyclic model for Iwr1-directed Pol II nuclear import (Figure 16): 1) Pol II is completely assembled in the cytoplasm (Boulon et al, 2010). 2) Iwr1 then binds Pol II (see 3.2.1) between the large Pol II subunits (see 3.3) maybe sensing complete assembly of the enzyme. 3) Next, Iwr1 uses its bipartite NLS (see 3.1.1) to direct Pol II nuclear import (see 3.1.3). 4) In the nucleus, Iwr1 is displaced from the Pol II active centre cleft during formation of the transcription initiation complex on promoter DNA (see 3.3.3). 5) Liberated Iwr1 is then exported from the nucleus with the help of its NES (see 3.1.1) by the karyopherin Xpo1 (Peiro-Chova & Estruch, 2009) for another round of Pol II nuclear import.

The data and model are consistent with Pol II assembly in the cytoplasm (Boulon et al, 2010) and Pol II disassembly and Rpb1 degradation in the nucleus (see 3.5.2). Iwr1 is specific for Pol II among the RNA polymerases, both in terms of binding and in its requirement for nuclear localization. Yet, similar mechanism might apply to Pol I and Pol III. In plants, Iwr1 may additionally direct nuclear import of the plant-specific Pol II-related enzymes Pol IV and Pol V (Ream et al, 2009), because Iwr1 is required for Pol IV/V-dependent processes in *Arabidopsis thaliana* (He et al, 2009; Kanno et al, 2010). Further, Iwr1 function seems to be conserved among all eukaryotes, including humans (see 3.1.4).

These results provide a starting point for analyzing the mechanisms underlying RNA polymerase biogenesis and subcellular localization, which may well constitute a mechanism within the recently suggested global transcriptional regulation (Zhurinsky et al, 2010).

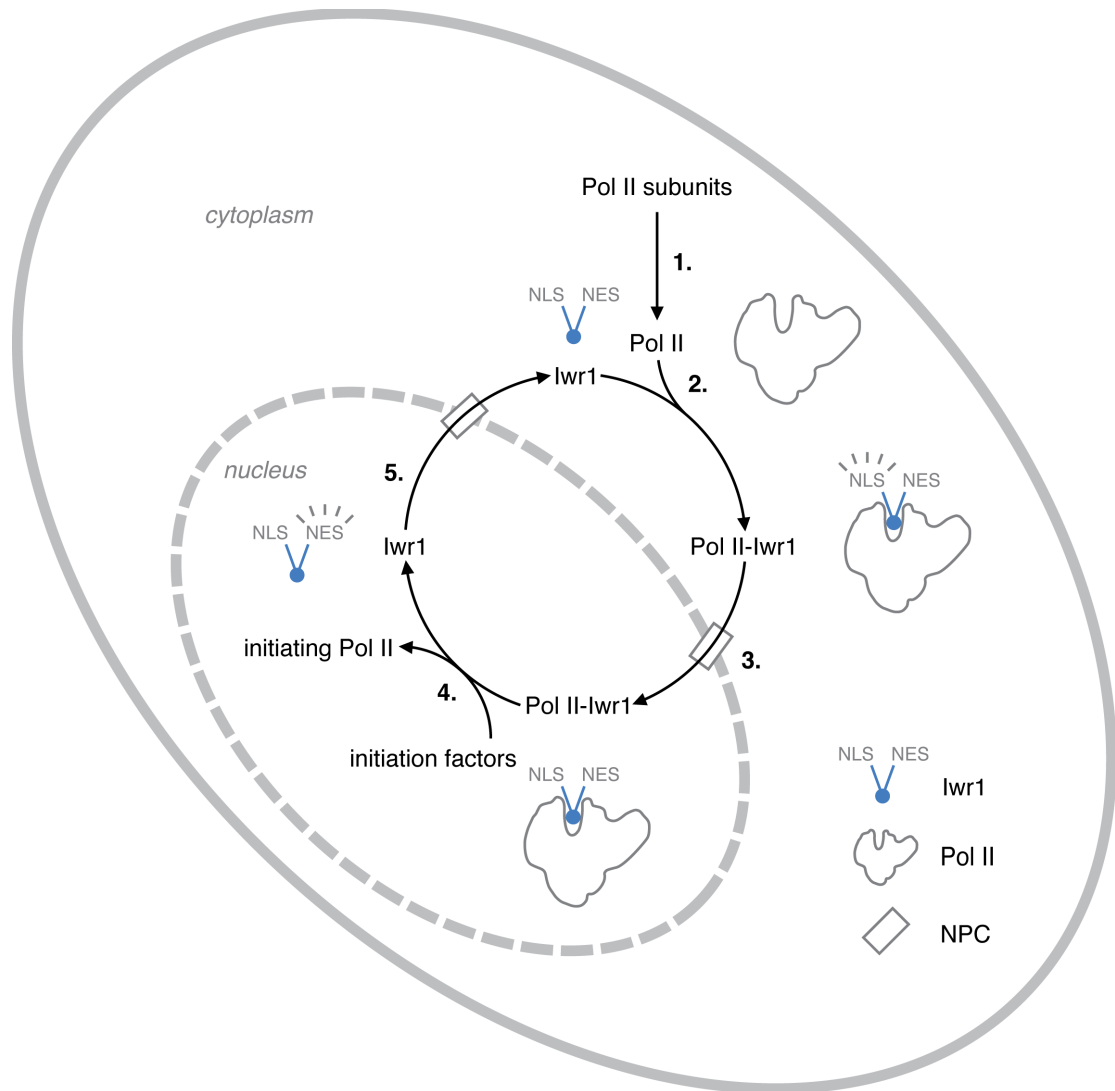


Figure 16 · Cyclic model for Iwr1-directed Pol II nuclear import. Pol II assembles from its subunits in the cytoplasm (1.). Iwr1 senses complete assembly by binding in the Pol II cleft (2.). The Iwr1 NLS directs Pol II import through the nuclear pore complex (NPC, 3.). Assembly of Pol II with initiation factors on promoter DNA releases Iwr1 (4.). The Iwr1 NES facilitates nuclear export of the factor, completing the cycle (5.).

4.2 Extending the proposed model

To further strengthen and extend the suggested model for Pol II nuclear import, I would like to suggest a series of experiments.

Interestingly, Iwr1 is not an essential protein in yeast. A sufficient amount of Pol II still translocates to the nucleus in $\Delta iwr1$ cells ensuring viability. However, essential nuclear factors often depend on import routes that are based on non-essential karyopherins (see 1.5.2). This seems to be rooted in the inherent redundancy in import pathways. Firstly, karyopherins overlap in their specificity for cargo. Secondly, functional interactions in the nucleus, such as a GTF binding to polymerase, might already take place in the nucleus, thereby providing fractional import of a partner that is orphaned in respect to his proper import route. Good candidates that might provide fractional Pol II import are TFIIB and TFIIF that both have a considerable affinity to Pol II and that possess a proper NLS. As the expression of non-functional Iwr1 NLS mutants resulted in slower growth than absence of the entire factor, binding of Iwr1 seems to interfere with one of the redundant import pathways. As TFIIF was co-purified together with Iwr1 (Krogan et al, 2006), TFIIB for which competition with Iwr1 was shown (see 3.3.3) seems to be the most likely candidate for providing redundant Pol II import. Therefore, if not lethal by themselves, TFIIB NLS mutations in the $\Delta iwr1$ background should be tested. In parallel, the accessibility of the TFIIB NLS for Kap114 in a Pol II-TFIIB complex should be examined *in vitro*.

Binding of Iwr1 in the cleft of Pol II suggests the role of an assembly sensor, ensuring nuclear import of only fully assembled complexes. In this context, the phenotype of an Iwr1 NLS fusion with either Rpb1 or Rpb3 in $\Delta iwr1$ cells would be of interest. Additionally providing NLS mutant Iwr1 to account for a possible role of Iwr1 in polymerase biogenesis (see 4.3) might provide further insights.

In the model it is assumed that the Iwr1 NLS is accessible in complex with Pol II and that the NES is not. In free Iwr1 the NES should be exposed and the NLS might be masked by some autoinhibitory mechanism. To clarify this, binding

of Kap60/Kap95, that recognize the classical bipartite NLS, and of Xpo1/RanGTP to free Iwr1 and Pol II-Iwr1 should be tested (see 1.5.1). In the case of Kap60, an N-terminal truncation without the autoinhibitory IBB domain could be used, opening up the possibility to omit Kap95 (Cook et al, 2007). Concerning Xpo1/RanGTP, a GTPase-deficient Ran mutation, such as Q69L is needed (Klebe et al, 1995).

To ultimately answer how Iwr1 interacts with Pol II and in which way localisation signals are masked or not, an atomic model of the Pol II-Iwr1 complex is needed. As a minimal and functional Pol II binding domain was delineated in the course of this work (see 3.2.2), co-crystallization with this domain or soaking of pre-formed Pol II crystals with even smaller variants that can still be expressed but do not bind with sufficient affinity for size-exclusion chromatography seems to be the best starting point. As the Rpb4/7 region in the cryo-EM of the Pol II-Iwr1 complex is shifted relatively to the position in the crystal structure (see 3.3.1) co-crystallization trials with 10-subunit core Pol II, lacking Rpb4/7, might be promising.

4.3 Iwr1 might integrate RNA polymerase II biogenesis and nuclear import

Recently, a role of HSP90 in Pol II biogenesis was published (Boulon et al, 2010). Interestingly, a protein co-purified with Iwr1 *in vivo* is Cns1 (Krogan et al, 2006), an essential HSP90 co-chaperone (Hainzl et al, 2004; Marsh et al, 1998; Tesic et al, 2003). In preliminary experiments during this work binding of recombinant Cns1 to Iwr1, but not to Pol II or Pol II-Iwr1 was observed (not shown). The most confident *in vivo* interaction of Iwr1 that was not a Pol II subunit is the kinetochore protein Cbf2 (Krogan et al, 2006). Cbf2 also undergoes interactions with HSP90 as a member of the CBF3 complex (Lingelbach & Kaplan, 2004; Pearl et al, 2008). Whatever might be the detailed nature of these interactions, they seem to establish a link between nuclear import and biogenesis of Pol II. A possible role might for example be in recruitment of the

two largest subunits to HSP90 or in picking up the fully assembled enzyme at the chaperone. Interestingly, Rpb1 protein levels and half-life are increased upon IWR1 deletion (see 3.5.2). Yet, according to ongoing experiments Rpb3 levels are unchanged, implying that a second function of Iwr1 should be specific for Rpb1 and possibly for Rpb2. To clarify this, a comprehensive comparison of all Pol II subunit levels between $\Delta iwr1$ and wild-type cells should be performed.

By approaching these questions, Iwr1 might emerge with dual roles and as a factor that integrates Pol II biogenesis and Pol II nuclear import.

5 References

Albertini M, Pemberton LF, Rosenblum JS, Blobel G (1998) A novel nuclear import pathway for the transcription factor TFIIIS. *J Cell Biol* **143**: 1447-1455

Albuquerque CP, Smolka MB, Payne SH, Bafna V, Eng J, Zhou H (2008) A multidimensional chromatography technology for in-depth phosphoproteome analysis. *Mol Cell Proteomics* **7**: 1389-1396

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

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

Beaudenon SL, Huacani MR, Wang G, McDonnell DP, Huibregtse JM (1999) Rsp5 ubiquitin-protein ligase mediates DNA damage-induced degradation of the large subunit of RNA polymerase II in *Saccharomyces cerevisiae*. *Mol Cell Biol* **19**: 6972-6979

Belle A, Tanay A, Bitincka L, Shamir R, O'Shea EK (2006) Quantification of protein half-lives in the budding yeast proteome. *Proc Natl Acad Sci U S A* **103**: 13004-13009

Bianchi MM, Ngo S, Vandenbol M, Sartori G, Morlupi A, Ricci C, Stefani S, Morlino GB, Hilger F, Carignani G, Slonimski PP, Frontali L (2001) Large-scale phenotypic analysis reveals identical contributions to cell functions of known and unknown yeast genes. *Yeast* **18**: 1397-1412

Biegert A, Mayer C, Remmert M, Soding J, Lupas AN (2006) The MPI Bioinformatics Toolkit for protein sequence analysis. *Nucleic Acids Res* **34**: W335-339

Biegert A, Soding J (2009) Sequence context-specific profiles for homology searching. *Proc Natl Acad Sci U S A* **106**: 3770-3775

Boulon S, Pradet-Balade B, Verheggen C, Molle D, Boireau S, Georgieva M, Azzag K, Robert M-C, Ahmad Y, Neel H, Lamond AI, Bertrand E (2010) HSP90 and its R2TP/Prefoldin-like co-chaperone are involved in the cytoplasmic assembly of RNA polymerase II. *Mol Cell*

Brachat S, Dietrich FS, Voegeli S, Zhang Z, Stuart L, Lerch A, Gates K, Gaffney T, Philippsen P (2003) Reinvestigation of the *Saccharomyces cerevisiae* genome annotation by comparison to the genome of a related fungus: *Ashbya gossypii*. *Genome Biol* **4**: R45

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248-254

Chen JZ, Grigorieff N (2007) SIGNATURE: a single-particle selection system for molecular electron microscopy. *J Struct Biol* **157**: 168-173

Chook YM, Blobel G (2001) Karyopherins and nuclear import. *Curr Opin Struct Biol* **11**: 703-715

Cloutier P, Coulombe B (2010) New insights into the biogenesis of nuclear RNA polymerases? *Biochem Cell Biol* **88**: 211-221

Collins SR, Miller KM, Maas NL, Roguev A, Fillingham J, Chu CS, Schuldiner M, Gebbia M, Recht J, Shales M, Ding H, Xu H, Han J, Ingvarsdottir K, Cheng B, Andrews B, Boone C, Berger SL, Hieter P, Zhang Z, Brown GW, Ingles CJ, Emili A, Allis CD, Toczyski DP, Weissman JS, Greenblatt JF, Krogan NJ (2007) Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. *Nature* **446**: 806-810

Cook A, Bono F, Jinek M, Conti E (2007) Structural biology of nucleocytoplasmic transport. *Annu Rev Biochem* **76**: 647-671

Coulombe B, Burton ZF (1999) DNA bending and wrapping around RNA polymerase: a "revolutionary" model describing transcriptional mechanisms. *Microbiol Mol Biol Rev* **63**: 457-478

Cramer P, Armache KJ, Baumli S, Benkert S, Brueckner F, Buchen C, Damsma GE, Dengl S, Geiger SR, Jasiak AJ, Jawhari A, Jennebach S, Kamenski T, Kettenberger H, Kuhn CD, Lehmann E, Leike K, Sydow JF, Vannini A (2008) Structure of eukaryotic RNA polymerases. *Annu Rev Biophys* **37**: 337-352

Cramer P, Bushnell DA, Fu J, Gnatt AL, Maier-Davis B, Thompson NE, Burgess RR, Edwards AM, David PR, Kornberg RD (2000) Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* **288**: 640-649

Cramer P, Bushnell DA, Kornberg RD (2001) Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292**: 1863-1876

Dalmay T, Hamilton A, Rudd S, Angell S, Baulcombe DC (2000) An RNA-dependent RNA polymerase gene in Arabidopsis is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. *Cell* **101**: 543-553

Damsma GE, Alt A, Brueckner F, Carell T, Cramer P (2007) Mechanism of transcriptional stalling at cisplatin-damaged DNA. *Nat Struct Mol Biol* **14**: 1127-1133

Dingwall C, Laskey RA (1991) Nuclear targeting sequences--a consensus? *Trends Biochem Sci* **16**: 478-481

Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**: 113

Flores O, Lu H, Killeen M, Greenblatt J, Burton ZF, Reinberg D (1991) The small subunit of transcription factor IIF recruits RNA polymerase II into the preinitiation complex. *Proc Natl Acad Sci U S A* **88**: 9999-10003

Flores O, Lu H, Reinberg D (1992) Factors involved in specific transcription by mammalian RNA polymerase II. Identification and characterization of factor IIF. *J Biol Chem* **267**: 2786-2793

Forget D, Lacombe AA, Cloutier P, Al-Khoury R, Bouchard A, Lavalley-Adam M, Faubert D, Jeronimo C, Blanchette M, Coulombe B (2010) The protein interaction network of the

human transcription machinery reveals a role for the GTPase RPAP4/GPN1 and microtubule assembly in nuclear import and biogenesis of RNA polymerase II. *Mol Cell Proteomics*

Fornerod M, Ohno M, Yoshida M, Mattaj IW (1997) CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell* **90**: 1051-1060

Frank J, Radermacher M, Penczek P, Zhu J, Li Y, Ladjadj M, Leith A (1996) SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* **116**: 190-199

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) *Protein Identification and Analysis Tools on the ExPASy Server*: Humana Press.

Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon AM, Cruciat CM, Remor M, Hofert C, Schelder M, Brajenovic M, Ruffner H, Merino A, Klein K, Hudak M, Dickson D, Rudi T, Gnau V, Bauch A, Bastuck S, Huhse B, Leutwein C, Heurtier MA, Copley RR, Edelman A, Querfurth E, Rybin V, Drewes G, Raida M, Bouwmeester T, Bork P, Seraphin B, Kuster B, Neubauer G, Superti-Furga G (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**: 141-147

Geiger SR, Lorenzen K, Schrieck A, Hanecker P, Kostrewa D, Heck AJ, Cramer P (2010) RNA polymerase I contains a TFIIF-related DNA-binding subcomplex. *Mol Cell* **39**: 583-594

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JY, Zhang J (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* **5**: R80

Goler-Baron V, Selitrennik M, Barkai O, Haimovich G, Lotan R, Choder M (2008) Transcription in the nucleus and mRNA decay in the cytoplasm are coupled processes. *Genes Dev* **22**: 2022-2027

Gouet P, Courcelle E, Stuart DI, Metoz F (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**: 305-308

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

Hainzl O, Wegele H, Richter K, Buchner J (2004) Cns1 is an activator of the Ssa1 ATPase activity. *J Biol Chem* **279**: 23267-23273

Harreman M, Taschner M, Sigurdsson S, Anindya R, Reid J, Somesh B, Kong SE, Banks CA, Conaway RC, Conaway JW, Svejstrup JQ (2009) Distinct ubiquitin ligases act sequentially for RNA polymerase II polyubiquitylation. *Proc Natl Acad Sci U S A* **106**: 20705-20710

He XJ, Hsu YF, Zhu S, Liu HL, Pontes O, Zhu J, Cui X, Wang CS, Zhu JK (2009) A conserved transcriptional regulator is required for RNA-directed DNA methylation and plant development. *Genes Dev*

Herr AJ, Jensen MB, Dalmay T, Baulcombe DC (2005) RNA polymerase IV directs silencing of endogenous DNA. *Science* **308**: 118-120

Hirata A, Murakami KS (2009) Archaeal RNA polymerase. *Curr Opin Struct Biol* **19**: 724-731

Hodges JL, Leslie JH, Mosammaparast N, Guo Y, Shabanowitz J, Hunt DF, Pemberton LF (2005) Nuclear import of TFIIB is mediated by Kap114p, a karyopherin with multiple cargo-binding domains. *Mol Biol Cell* **16**: 3200-3210

Hu Z, Killion PJ, Iyer VR (2007) Genetic reconstruction of a functional transcriptional regulatory network. *Nat Genet* **39**: 683-687

Huh WK, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, O'Shea EK (2003) Global analysis of protein localization in budding yeast. *Nature* **425**: 686-691

Huibregtse JM, Yang JC, Beaudenon SL (1997) The large subunit of RNA polymerase II is a substrate of the Rsp5 ubiquitin-protein ligase. *Proc Natl Acad Sci U S A* **94**: 3656-3661

Ishiguro A, Kimura M, Yasui K, Iwata A, Ueda S, Ishihama A (1998) Two large subunits of the fission yeast RNA polymerase II provide platforms for the assembly of small subunits. *J Mol Biol* **279**: 703-712

Ishihama A (1981) Subunit of assembly of Escherichia coli RNA polymerase. *Adv Biophys* **14**: 1-35

Ishihama A, Ito K (1972) Subunits of RNA polymerase in function and structure. II. Reconstitution of Escherichia coli RNA polymerase from isolated subunits. *J Mol Biol* **72**: 111-123

Janke C, Magiera MM, Rathfelder N, Taxis C, Reber S, Maekawa H, Moreno-Borchart A, Doenges G, Schwob E, Schiebel E, Knop M (2004) A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast* **21**: 947-962

Jones M. (2001) Isolation of plasmid DNA from yeast using the QIAprep® Spin Miniprep Kit. *Qiagen User-Developed Protocol*.

Juneau K, Palm C, Miranda M, Davis RW (2007) High-density yeast-tiling array reveals previously undiscovered introns and extensive regulation of meiotic splicing. *Proc Natl Acad Sci U S A* **104**: 1522-1527

Kanno T, Bucher E, Daxinger L, Huettel B, Kreil DP, Breinig F, Lind M, Schmitt MJ, Simon SA, Gurazada SG, Meyers BC, Lorkovic ZJ, Matzke AJ, Matzke M (2010) RNA-directed DNA methylation and plant development require an IWR1-type transcription factor. *EMBO Rep* **11**: 65-71

Kanno T, Huettel B, Mette MF, Aufsatz W, Jaligot E, Daxinger L, Kreil DP, Matzke M, Matzke AJ (2005) Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nat Genet* **37**: 761-765

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

Kettenberger H, Eisenfuhr A, Brueckner F, Theis M, Famulok M, Cramer P (2006) Structure of an RNA polymerase II-RNA inhibitor complex elucidates transcription regulation by noncoding RNAs. *Nat Struct Mol Biol* **13**: 44-48

Kimura M, Ishiguro A, Ishihama A (1997) RNA polymerase II subunits 2, 3, and 11 form a core subassembly with DNA binding activity. *J Biol Chem* **272**: 25851-25855

Kimura M, Ishihama A (2000) Involvement of multiple subunit-subunit contacts in the assembly of RNA polymerase II. *Nucleic Acids Res* **28**: 952-959

Kimura M, Sakurai H, Ishihama A (2001) Intracellular contents and assembly states of all 12 subunits of the RNA polymerase II in the fission yeast *Schizosaccharomyces pombe*. *Eur J Biochem* **268**: 612-619

Klebe C, Bischoff FR, Ponstingl H, Wittinghofer A (1995) Interaction of the nuclear GTP-binding protein Ran with its regulatory proteins RCC1 and RanGAP1. *Biochemistry* **34**: 639-647

Kolodziej PA, Young RA (1991) Mutations in the three largest subunits of yeast RNA polymerase II that affect enzyme assembly. *Mol Cell Biol* **11**: 4669-4678

Kostrewa D, Zeller ME, Armache K-J, Seizl M, Leike K, Thomm M, Cramer P (2009) RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature* **462**: 323-330

Kosugi S, Hasebe M, Matsumura N, Takashima H, Miyamoto-Sato E, Tomita M, Yanagawa H (2009a) Six classes of nuclear localization signals specific to different binding grooves of importin alpha. *J Biol Chem* **284**: 478-485

Kosugi S, Hasebe M, Tomita M, Yanagawa H (2009b) Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proc Natl Acad Sci U S A* **106**: 10171-10176

Krogan NJ, Cagney G, Yu H, Zhong G, Guo X, Ignatchenko A, Li J, Pu S, Datta N, Tikuisis AP, Punna T, Peregrin-Alvarez JM, Shales M, Zhang X, Davey M, Robinson MD, Paccanaro A, Bray JE, Sheung A, Beattie B, Richards DP, Canadien V, Lalev A, Mena F, Wong P,

Starostine A, Canete MM, Vlasblom J, Wu S, Orsi C, Collins SR, Chandran S, Haw R, Rilstone JJ, Gandhi K, Thompson NJ, Musso G, St Onge P, Ghanny S, Lam MH, Butland G, Altaf-Ul AM, Kanaya S, Shilatifard A, O'Shea E, Weissman JS, Ingles CJ, Hughes TR, Parkinson J, Gerstein M, Wodak SJ, Emili A, Greenblatt JF (2006) Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* **440**: 637-643

la Cour T, Kiemer L, Molgaard A, Gupta R, Skriver K, Brunak S (2004) Analysis and prediction of leucine-rich nuclear export signals. *Protein Eng Des Sel* **17**: 527-536

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680-685

Lalo D, Carles C, Sentenac A, Thuriaux P (1993) Interactions between three common subunits of yeast RNA polymerases I and III. *Proc Natl Acad Sci U S A* **90**: 5524-5528

Lingelbach LB, Kaplan KB (2004) The interaction between Sgt1p and Skp1p is regulated by HSP90 chaperones and is required for proper CBF3 assembly. *Mol Cell Biol* **24**: 8938-8950

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

Marsh JA, Kalton HM, Gaber RF (1998) Cns1 is an essential protein associated with the hsp90 chaperone complex in *Saccharomyces cerevisiae* that can restore cyclophilin 40-dependent functions in *cpr7*Delta cells. *Mol Cell Biol* **18**: 7353-7359

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

Morehouse H, Buratowski RM, Silver PA, Buratowski S (1999) The importin/karyopherin Kap114 mediates the nuclear import of TATA-binding protein. *Proc Natl Acad Sci U S A* **96**: 12542-12547

Mosammaparast N, Pemberton LF (2004) Karyopherins: from nuclear-transport mediators to nuclear-function regulators. *Trends Cell Biol* **14**: 547-556

Needleman SB, Wunsch CD (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol* **48**: 443-453

Nguyen VT, Giannoni F, Dubois MF, Seo SJ, Vigneron M, Kedinger C, Bensaude O (1996) In vivo degradation of RNA polymerase II largest subunit triggered by alpha-amanitin. *Nucleic Acids Res* **24**: 2924-2929

Niall HD (1973) Automated Edman degradation: the protein sequenator. *Methods Enzymol* **27**: 942-1010

Nonet M, Scafe C, Sexton J, Young R (1987) Eucaryotic RNA polymerase conditional mutant that rapidly ceases mRNA synthesis. *Mol Cell Biol* **7**: 1602-1611

Page N, Gerard-Vincent M, Menard P, Beaulieu M, Azuma M, Dijkgraaf GJ, Li H, Marcoux J, Nguyen T, Dowse T, Sdicu AM, Bussey H (2003) A *Saccharomyces cerevisiae* genome-wide mutant screen for altered sensitivity to K1 killer toxin. *Genetics* **163**: 875-894

Pearl LH, Prodromou C, Workman P (2008) The Hsp90 molecular chaperone: an open and shut case for treatment. *Biochem J* **410**: 439-453

Peiro-Chova L, Estruch F (2009) The yeast RNA polymerase II-associated factor Iwr1p is involved in the basal and regulated transcription of specific genes. *J Biol Chem* **284**: 28958-28967

Pemberton LF, Paschal BM (2005) Mechanisms of receptor-mediated nuclear import and nuclear export. *Traffic* **6**: 187-198

Pollard VW, Michael WM, Nakielny S, Siomi MC, Wang F, Dreyfuss G (1996) A novel receptor-mediated nuclear protein import pathway. *Cell* **86**: 985-994

Pouton CW, Wagstaff KM, Roth DM, Moseley GW, Jans DA (2007) Targeted delivery to the nucleus. *Adv Drug Deliv Rev* **59**: 698-717

- Rani PG, Ranish JA, 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 JA, Hahn S (1991) The yeast general transcription factor TFIIA is composed of two polypeptide subunits. *J Biol Chem* **266**: 19320-19327
- Ream TS, Haag JR, Wierzbicki AT, Nicora CD, Norbeck AD, Zhu JK, Hagen G, Guilfoyle TJ, Pasa-Tolic L, Pikaard CS (2009) Subunit compositions of the RNA-silencing enzymes Pol IV and Pol V reveal their origins as specialized forms of RNA polymerase II. *Mol Cell* **33**: 192-203
- Reid J, Svejstrup JQ (2004) DNA damage-induced Def1-RNA polymerase II interaction and Def1 requirement for polymerase ubiquitylation in vitro. *J Biol Chem* **279**: 29875-29878
- Reimand J, Vaquerizas JM, Todd AE, Vilo J, Luscombe NM (2010) Comprehensive reanalysis of transcription factor knockout expression data in *Saccharomyces cerevisiae* reveals many new targets. *Nucleic Acids Res*
- Ribar B, Prakash L, Prakash S (2007) ELA1 and CUL3 are required along with ELC1 for RNA polymerase II polyubiquitylation and degradation in DNA-damaged yeast cells. *Mol Cell Biol* **27**: 3211-3216
- Rosenblum JS, Pemberton LF, Bonifaci N, Blobel G (1998) Nuclear import and the evolution of a multifunctional RNA-binding protein. *J Cell Biol* **143**: 887-899
- Rubbi L, Labarre-Mariotte S, Chedin S, Thuriaux P (1999) Functional characterization of ABC10alpha, an essential polypeptide shared by all three forms of eukaryotic DNA-dependent RNA polymerases. *J Biol Chem* **274**: 31485-31492
- Sakurai H, Ishihama A (2002) Level of the RNA polymerase II in the fission yeast stays constant but phosphorylation of its carboxyl terminal domain varies depending on the phase and rate of cell growth. *Genes Cells* **7**: 273-284

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

Selth LA, Sigurdsson S, Svejstrup JQ (2010) Transcript Elongation by RNA Polymerase II. *Annu Rev Biochem* **79**: 271-293

Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* **3**: Article3

Soding J, Biegert A, Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* **33**: W244-248

Somesh BP, Reid J, Liu WF, Sogaard TM, Erdjument-Bromage H, Tempst P, Svejstrup JQ (2005) Multiple mechanisms confining RNA polymerase II ubiquitylation to polymerases undergoing transcriptional arrest. *Cell* **121**: 913-923

Somesh BP, Sigurdsson S, Saeki H, Erdjument-Bromage H, Tempst P, Svejstrup JQ (2007) Communication between distant sites in RNA polymerase II through ubiquitylation factors and the polymerase CTD. *Cell* **129**: 57-68

Stade K, Ford CS, Guthrie C, Weis K (1997) Exportin 1 (Crm1p) is an essential nuclear export factor. *Cell* **90**: 1041-1050

Suel KE, Chook YM (2009) Kap104p imports the PY-NLS-containing transcription factor Tfg2p into the nucleus. *J Biol Chem* **284**: 15416-15424

Sun ZW, Hampsey M (1995) Identification of the gene (SSU71/TFG1) encoding the largest subunit of transcription factor TFIIF as a suppressor of a TFIIB mutation in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* **92**: 3127-3131

Tesic M, Marsh JA, Cullinan SB, Gaber RF (2003) Functional interactions between Hsp90 and the co-chaperones Cns1 and Cpr7 in *Saccharomyces cerevisiae*. *J Biol Chem* **278**: 32692-32701

Thomas MC, Chiang CM (2006) The general transcription machinery and general cofactors. *Crit Rev Biochem Mol Biol* **41**: 105-178

Titov AA, Blobel G (1999) The karyopherin Kap122p/Pdr6p imports both subunits of the transcription factor IIA into the nucleus. *J Cell Biol* **147**: 235-246

Wayne Davis M. (2009) ApE- A Plasmid Editor

Woudstra EC, Gilbert C, Fellows J, Jansen L, Brouwer J, Erdjument-Bromage H, Tempst P, Svejstrup JQ (2002) A Rad26-Def1 complex coordinates repair and RNA pol II proteolysis in response to DNA damage. *Nature* **415**: 929-933

Wu Z, Irizarry RA (2004) Preprocessing of oligonucleotide array data. *Nat Biotechnol* **22**: 656-658; author reply 658

Yan Q, Moreland RJ, Conaway JW, Conaway RC (1999) Dual roles for transcription factor IIF in promoter escape by RNA polymerase II. *J Biol Chem* **274**: 35668-35675

Yasui K, Ishiguro A, Ishihama A (1998) Location of subunit-subunit contact sites on RNA polymerase II subunit 3 from the fission yeast *Schizosaccharomyces pombe*. *Biochemistry* **37**: 5542-5548

Zecherle GN, Whelen S, Hall BD (1996) Purines are required at the 5' ends of newly initiated RNAs for optimal RNA polymerase III gene expression. *Mol Cell Biol* **16**: 5801-5810

Zhurinsky J, Leonhard K, Watt S, Marguerat S, Bahler J, Nurse P (2010) A Coordinated Global Control over Cellular Transcription. *Curr Biol*

Abbreviations

a. a.	amino acids
bp	base pairs
C-terminus	carboxy terminus
CAPS	N-cyclohexyl-3-aminopropanesulfonic acid
ChIP	chromatin immunoprecipitation
CTD	carboxy-terminal domain
CV	column volumes
Da	Dalton
DAPI	4',6-diamidino-2-phenylindole
DIC	differential interference contrast
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTT	1,4-dithio-D,L-threitol
EGFP	enhanced green fluorescent protein
EM	electron microscopy
et al.	et alii (Latin "and others")
g	earth's gravity
GFP	green fluorescent protein
GTF	general transcription factor
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethane sulfonic acid
HSP	heat shock protein
IMAC	immobilized metal ion affinity chromatography
IPTG	isopropyl- β -D-thiogalactopyranoside
Iwr1	Interact With RNA polymerase II
kDa	kilo Dalton
LB	Luria-Bertani (media)
M	molar
MOPS	4-morpholinepropanesulfonic acid
mRNA	messenger ribonucleic acid
MS	mass spectrometry
N-terminus	amino terminus
NES	nuclear export sequence
NLS	nuclear localization sequence
NPC	nuclear pore complex
NTA	nitrilotriacetic acid

OD_{x nm}	optical density at a wavelength of X nm
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	protein data bank
PI	protease inhibitor
PIC	pre-initiation complex
Pol	RNA polymerase
PVDF	polyvinylidene fluoride
RNA	ribonucleic acid
rpm	rounds per minute
RT-PCR	reverse transcriptase PCR
S	Svedberg (sedimentation coefficient unit)
SDS	sodium dodecylsulfate
SGA	synthetic genetic array
TAF	TBP-associated factor
TAP	tandem affinity purification
TCA	trichloroacetic acid
TCEP	tris(2-carboxyethyl)phosphine
TCR	transcription-coupled repair
TF	transcription factor
Tris	tris-(hydroxymethyl)-aminomethane
U	units
URA	uracil
UTR	untranslated region
v/v	volume per volume
w/v	weight per volume
YPD	yeast extract peptone dextrose

Curriculum Vitae

Elmar Czeko · Biochemist (Diploma)

born 23rd of December, 1980 in Temeschburg, Romania

Nationality German · Family status: married, one child

RESEARCH

2007.01 – 2011.02 Doctoral thesis at the Gene Center of the Ludwig-Maximilians-Universität, Munich with Patrick Cramer (*Structural and Functional Investigation of RNA Polymerase II Nuclear Import*)

EDUCATION

2006.01 – 2006.07 Diploma thesis at The Rockefeller University, New York with James Darnell (*Examining the Cooperative Binding between STAT3 and AP-1 on DNA*)

2001.10 – 2006.07 Studies of Biochemistry (Diploma course) at the Freie Universität Berlin with research stays in Frankfurt am Main, Paris and London

2000.05 Abitur (general qualification for university entrance),
Gymnasium Ernestinum Coburg

SCHOLARSHIPS

2007.05 – 2009.04 Kekulé PhD Scholarship of the Fund of the German Chemical Industry (FCI)

2006.01 – 2006.05 Scholarship of the German Academic Exchange Service (DAAD)

2001.10 – 2006.03 Bayerisches Begabtenstipendium (Bavarian University Scholarship)

PUBLICATIONS

- Czeko E, Seizl M, Augsberger C, Mielke T, Cramer P (2011). Iwr1 directs RNA polymerase II nuclear import. *Mol Cell*.
- Ginsberg M, Czeko E, Muller P, Ren Z, Chen X, Darnell JE (2007). Amino acid residues required for physical and cooperative transcriptional interaction of STAT3 and AP-1 proteins c-Jun and c-Fos. *Mol Cell Biol* 27, 6300-6308.

SCIENTIFIC CONFERENCES AND SEMINARS

- 2010.10 *Annual conference of the Center for Integrated Protein Sciences (Presentation)*
- 2010.08 *EMBO Conference Transcription and chromatin (Poster presentation)*
- 2010.06 *EMBO Conference Gene transcription in yeast (Poster presentation)*
- 2010.03 *Life science PhD student symposium 2010, munich-interact.org (Organizer)*
- 2009.11 *Workshop on advanced strategies for protein and protein complex expression in yeast, Institute de Biologia Molecular de Barcelona (Presentation)*
- 2009.07 *59th Nobel Laureate Meeting at Lake Constance, Stiftung Lindauer Nobelpreisträgertreffen (Participant)*
- 2009.04 *Life science PhD student symposium 2009, munich-interact.org (Speaker of the Organization Committee)*
- 2008.08 *EMBO-Course Cryo-electron microscopy and 3D-image analysis (Participant, Presentation)*