Functional Characterization of *S. cerevisiae* Upstream Activating Factor and its subunits

DISSERTATION ZUR ERLANGUNG EINES DOKTORGRADES DER NATURWISSENSCHAFTEN (DR. RER. NAT.)

FAKULTÄT FÜR BIOCHEMIE UND VORKLINISCHE MEDIZIN

UNIVERSITÄT REGENSBURG

vorgelegt von:

Philipp Becker

aus Willich

September 2023

Die Arbeit wurde angeleitet von: Prof. Dr. Christoph Engel

1. Gutachter	Prof. Dr. Christoph Engel
2. Gutachter	Prof. Dr. Herbert Tschochner

Verteidigung

Drittprüfer:	Prof. Dr. Remco Sprangers
Ersatzprüfer:	Prof. Dr. Gunter Meister
Prüfungsvorsitz:	Prof. Dr. Joachim Griesenbeck

Promotionsgesuch eingereicht am: 14.09.2023

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt, dass ich die vorgelegte Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus den anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe des Literaturzitats gekennzeichnet.

_____

Philipp Becker Regensburg, 14.09.2023

Contents

Contributions

Many great people assisted with their knowledge and know-how regarding different aspects of this project. With this, I want to thank all of you for your contributions. Your help was very much appreciated.

Thanks to …

Christoph Engel: who supervised and provided the idea and the general direction for the project. Further, he partly provided expression vectors and purification protocols for a quick start of the project. Most importantly, advice was given whenever necessary.

Michael Pilsl: who provided useful UAF knowledge, protocols, promoter templates, UAF purified from insect cells, and Pol I and its transcription factors for in vitro transcription assays.

Herbert Tschochner who, as part of the UAF team, helped with feedback and provided new ideas.

Joachim Griesenbeck who was always available for discussions and a very helpful and critical voice regarding experiment design, controls, etc.

Astrid Bruckmann who helped with mass spectrometry analysis of several samples.

Mona Höcherl: who helped with preparation of proteins used in this project.

Antonia Neumeier: who helped with the cloning of UAF co-expression vectors and UAF expression/purification during her internship.

David Stelzig: who helped with fluorescence anisotropy.

Maximilian Pichler: who did the statistical evaluation of the promoter fragment EMSAs.

Flo & Julia for many lab-related discussions and advice.

1 Summary

In the baker's yeast '*Saccharomyces cerevisiae*', production of all ribosomal RNAs, except one, relies on transcription by RNA-Polymerase I (Pol I). Transcription of the rDNA gene follows the formation of the Pol I pre-initiation complex (Pol I PIC). Over the last decades, numerous studies have provided insight into structure and function of Pol I, as well as its transcription factors Rrn3, TATA-binding protein (TBP), and Core Factor (CF). The fourth transcription factor – the upstream activating factor 'UAF' – has been shown to function as a potent enhancer of Pol I transcription initiation, repressor of RNA-Pol-II transcription, and as a maintainer of a native rDNA copy levels. However, to the day this project was started, structural information about this factor was completely missing. With this project, we aimed at providing additional functional insight and structural information about this important complex in order to contribute to a more comprehensive understanding of transcription initiation by Pol I.

During this project, expression and purification protocols for UAF subunits and the whole complex were optimized, followed by functional characterization and crystallization attempts. Functional characterization of subunits Uaf30 and the Rrn9 / Rrn10 subcomplex showed, that both – Uaf30 and Rrn9 – play an important role for the interaction with the upstream element. We show that promoterbinding by Rrn9 is likely unspecific and that it relies on a C terminal DNA binding domain. Uaf30, in contrast, shows a significant preference towards an upstream region of the upstream element around position -100 relative to the transcription start site. From this data, we hypothesize, that promoter targeting by UAF may rely on tethering by Uaf30, while Rrn9 cooperatively binds the promoter further downstream in an unspecific manner to strengthen the interaction between UAF and DNA.

To challenge these claims, Rrn9-C and Uaf30 deletion mutants were additionally tested in context of the whole UAF complex. In line with the previous results, deletion of Uaf30 leads to a notable loss of specificity, while deletion of Rrn9-C only exhibits reduced overall affinity. A double deletion mutant combines both loss of affinity and specificity. The impact of both deletions on the stimulation of transcription initiation by Pol I was tested via in vitro transcription assays. Interestingly, deletion of Uaf30 almost completely abolishes the transcription enhancement by UAF when compared to the wildtype complex and ΔC Rrn9 mutant. Contrary to expectations, in the latter, while transcription of the wildtype DNA template is less enhanced compared to the wildtype complex, transcription of the control template without an upstream element was significantly enhanced.

With the aim of providing structural data, a variety of constructs were subjected to high throughput crystallization screenings. Although isolation of crystallization grade protein was achieved for many of these constructs, including homologs from the baker's yeasts thermophilic relative - *Chaetomium thermophilum*, the bottleneck of actual protein crystallization turned out to be a major limiting factor. Only very few crystals were obtained and measured at a synchrotron. Neither did these crystals diffract, nor could they be reproduced.

2 Introduction

2.1 The rDNA genes

The need for protein synthesis is a common trait that is shared between all living organisms. Translation of mRNA into polypeptides requires ribosomes, which can be found across all species. While ribosomes of procaryotes and eucaryotes differ significantly in size, composition, and key mechanisms like translation initiation, termination, and regulation, their core structure is conserved, and is made possible by the structural and functional interplay of ribosomal proteins (RPs) and ribosomal RNAs (rRNAs) (Ben-Shem et al. 2011; Hinnebusch 2014; Jenner et al. 2012; Schmeing et al. 2009; Sonenberg et al. 2009). In eucaryotes, like the budding yeast *Saccharomyces cerevisiae*, the large (60 S) subunit of a ribosome consists of 46 ribosomal proteins and three rRNAs, the 5 S, the 5.8 S, and the 28 S rRNA. The small subunit (40 S) is made up of 33 RPs and the 18 S rRNA (Ben-Shem et al. 2011; Jenner et al. 2012; Woolford and Baserga 2013). The immense amounts of ribosomes that are necessary for maintaining native levels of mRNA translation during all stages of the cell cycle, makes ribosomal RNA the most abundand RNA (up to 80%) of all cellular RNAs (Laferté et al. 2006; Warner JR 1999). For reference: independently of the cell cycle, *S. cerevisiae* keeps a constant ratio of 10 ribosomes for each mRNA (Rudra et al. 2004; Zhao et al. 2003) and produces roughly 2000 ribosomes each minute (Warner JR 1999). This renders the production of ribosomal RNAs for ribosome biogenesis one of the most important, complex, and tightly regulated processes in all organisms, which, in case of misregulation, is linked to risk of cancer (Drygin et al. 2010; Montanaro et al. 2008).

While transcription of the shortest (5 S) rRNA is done by the RNA polymerase III, which is primarily responsible for synthesis of tRNA and other structured noncoding RNAs (Kharde et al. 2015; Kressler et al. 2012; Nikitina and Tishchenko 2005; White 2004), the 5.8 S, 18 S, and 25 S rRNA are transcribed by a dedicated enzyme, the RNA-Polymerase I (from here on referenced as **Pol I**) (Engel et al. 2017; Vannini and Cramer 2012). This 14 – subunit complex specifically transcribes rDNA genes, which, in *S. cerevisiae*, are located in one cluster on chromosome VII, which comprises roughly 150 repeats of the rDNA gene. Transcription of the rDNA genes yields the 35 S rRNA precursor (yeast) and leads to formation of the nucleolus, providing a cellular compartment for ribosome biogenesis and maturation. While, in the budding yeast, all rDNA gene repeats are clustered in one locus, resulting in a single

nucleolus, more than one of these nucleolar organization regions (NORs) can exist, corresponding to the number of rDNA loci in other species (e.g. 5 in humans) (Henderson et al. 1972; McClintock 1934; Moss et al. 2007; Rowe et al. 1996). The nucleolus forms substructures called fibrillar centers, areas where co-transcriptional assembly of the ribosomal subunits takes place (Cheutin et al. 2002; Scheer et al. 1993; Shaw and Doonan 2005; Trumtel et al. 2000), and outer granular centers, where the 40 S and the 60 S subunit are matured independently (Léger-Silvestre et al. 1999)(Cmarko et al. 2008). These highly complex processes of co-transcriptional processing and maturation involve a distinct pattern of cleavages and specific RNA modifications, which are dependent on a vast number of small nucleolar RNAs (snoRNAs) and other factors, which will not be further addressed in this thesis (Dörner et al. 2023; Granneman and Baserga 2005)(Cerezo et al. 2019; Ferreira-Cerca et al. 2005; Fromont-Racine et al. 2003; Milkereit et al. 2001; Thomson et al. 2013).

Unlike rRNA, RPs are synthesized outside of the nucleoli, and thus have to be imported for ribosome assembly. While this is usually also true for the 5 S rRNA (Moss et al. 2007; Németh 2010), in *S. cerevisiae* the 5 S gene is located in the intergenic spacer (IGS) of the rDNA gene and is hence transcribed inside the nucleolus (Moss et al. 2007; Thompson et al. 2003) *(Figure 1 B)*. Besides the sequence that actually codes for the rRNA precursor, every gene repeat harbors additional elements like the noncoding intergenic spacers or non transcribed spacers (NTS1 and NTS2, also calles IGS), repeating enhancers, termination sites, E-pros, and the rDNA promoter (Grummt et al. 1986; S. Henderson and Sollner-Webb 1986; Moss et al. 2007; Nikolaev et al. 1979; Srivastava and Schlessinger 1991). These elements have regulatory functions, controlling the transcription and replication of the rDNA (Elion and Warner 1986; Elaine et al. 1984; Kulkens et al. 1991; Musters et al. 1989) and help maintain the size of the rDNA locus (Hori et al. 2023; Kobayashi 2011; Kobayashi and Ganley 2005).

Although the sequence of the rDNA promoter itself is not really conserved among species, it usually comprises two elements, the upstream element (UE), which in yeast is the binding site of the Upstream Activating Factor (UAF), and the core element (CE) which is targeted by the Core Factor (CF) (Sollner-Webb et al. 1986; Moss et al. 2007; Moss et al. 1985)(Boukhgalter et al. 2002). Both elements are located directly upstream of the transcription start site (TSS) and are crucial for Pol I transcription initiation. Position, as well as spacing between both elements is essential for efficient preinitiation complex formation and initiation (Bordi et al. 2001; Choe, Schultz, and Reeder 1992; Kulkens et al. 1991; Musters

et al. 1989; Pikaard et al. 1990). The associated transcription factors CF and UAF will be discussed in more detail in the next chapters.

Since ribosome homeostasis is a pivotal element of cell fitness and survival, consequently, Pol I transcription is of equivalent importance. The exacting cellular demands for rRNA are met by high level rDNA transcription rates, which can only be achieved by simultaneous loading of multiple Pol I molecules onto each actively transcribed rDNA gene. With up to 100+ Pol I molecules per gene, loading rates are vastly increased compared to the Pol II and III system (French et al. 2003; Schneider et al. 2007). The highly optimized structure of Pol I, together with its transcription initiation factors, facilitates this by swift promoter melting, promoter escape, and its fast elongation speed (Patrick Cramer et al. 2000; Engel et al. 2013, 2017; Pilsl and Engel 2020). The loading of multiple Pol I molecules and nascent rRNA, as well as the co-transcriptional assembly of pre – ribosomal intermediates was first visualized in 1969 with an electron microscopy image of an rDNA gene fully occupied by Pol I, showing also the nascent rRNA precursors and pre – ribosomal particles, all together resembling a Christmas tree like structure, also known as 'Miller spread' *(Figure 1 A)* (Miller and Beatty 1969).

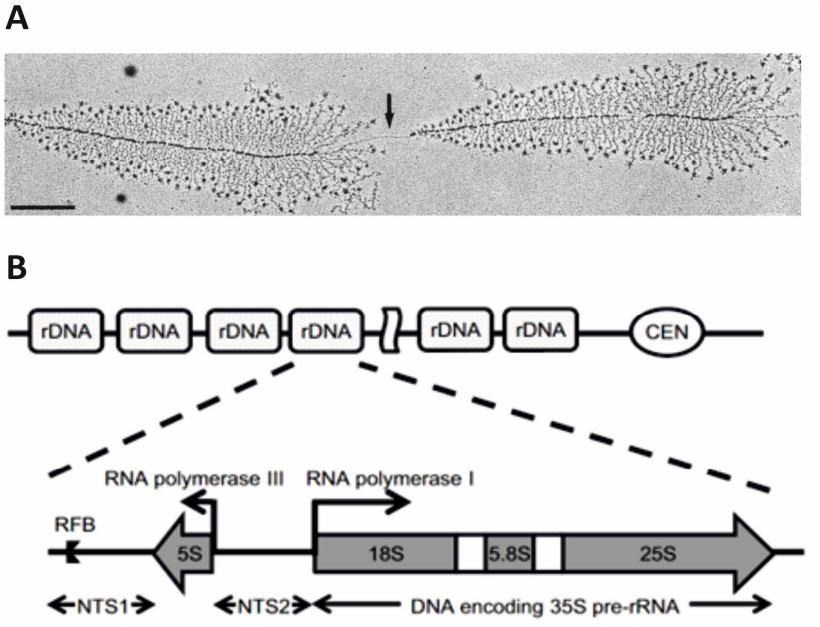


Figure 1: rDNA gene organization in S. cerevisiae.

A) Electron microscopy image of a Miller Spread, also known as Christmas Tree structure. Two transcribed genes loaded with polymerases and nascent pre-ribosomes at the ends of each branch. Genes are linked with intergenic spacers (black arrow) (Miller and Beatty 1969). **B)** Schematic architecture of the rDNA gene locus harboring the repetitive coding sequence for the 35S rDNA, the 5S rDNA, intergenic or non transcribed spacers (NTS1+2), and the replication fork barrier (RFB) (Sasano et al. 2017).

2.2 The Pol I transcription cycle

The transcription cycle of Pol I essentially includes three steps: Initiation, elongation, and termination. The following chapters will explore transcription initiation in more detail. However, elongation and termination of Pol I transcription are equally important and are also tightly regulated processes, aided by several factors that ensure accurate rRNA production. While a detailed review of both, elongation, and termination, would go beyond the scope of this thesis, the following simplified overview of the involved steps and factors will provide context for better understanding. For a more detailed review see: (Cramer 2019; Marques et al. 2013)

Pol I has been shown to be able to regulate its availability through hibernation in inactive dimers *(Figure 2)*, e.g. under nutrient starvation conditions (Engel et al. 2018; Fernández-Tornero 2018; Torreira et al. 2017). Recently, this was also observed in *Schizosaccharomyces pombe* (Heiss et al. 2021). This state facilitates response to reduced demands for ribosome biogenesis. As previously explained, Rrn3 render monomeric Pol I active, and ready for recruitment by the core factor. The assembly of the Pol I pre-initiation complex (PIC) will be explained in more detail in the next chapter.

Once the Pol I PIC is fully assembled to the transcription start site, the DNA strand is opened (promoter melting) and transcription of the 35 S rRNA can commence (transcription initiation). During this process, the Polymerase cleft changes its confirmation to a contracted state, tightly wrapping around the single stranded rDNA (Engel et al. 2017; Pilsl and Engel 2020; Sadian et al. 2019). When switching into elongation mode, transcription initiation factors (UAF / TBP / CF / Rrn3) are left behind (promoter escape) and elongation factors like Spt4/5, Spt6, the Paf1 complex, or Hmo1 take over to ensure processivity of Pol I (Anderson et al. 2011; Ucuncuoglu et al. 2016; Viktorovskaya et al. 2011; Zhang et al. 2009). Further, Pol I subunit A34.5 and A49 shows structural and functional similarity to the elongation factor TFIIF of the Pol II system, and subunit A12.2 facilitates RNA cleavage that is done by TFIIS (Engel et al. 2017; Geiger and Geiger 2010; Haag and Pikaard 2007; Kuhn et al. 2007). Pol I proceeds to elongate the nascent pre-rRNA while the RNA co-transcriptionally associates with components of the processing machinery, which is responsible for the highly complex maturation of the rRNA and assembly of ribosomes, which has been extensively reviewed:

(Cerezo et al. 2019; Dörner et al. 2023; Granneman and Baserga 2005; Klinge and Woolford 2019; Woolford and Baserga 2013).

After the transcript has been fully elongated, transcription ends. The exact mechanism behind transcription termination is not yet fully understood. However, a T rich sequence has been identified as termination site for Pol I, followed by a Reb1 binding site. Reb1, the *S. cerevisiae* homolog of transcription termination factor 1 (TTF1) in mammals, was demonstrated to terminate Pol I transcription I vitro. (Längst et al. 1998; Merkl et al. 2014; Merkl 2013; Reiter et al. 2012)(Mason et al. 1997). Interaction with the transcription release factor PTRF has been shown in mammals and yeast (Jansa and Grummt 1999), suggesting a conserved mechanism involving TTF1 / Reb1 and PTRF. Other models suggest e.g. the torpedo model, which is also found in Pol II termination. Here, the combined activity of Rnt1 (endonuclease), Rat1 (exonuclease), and Sen1 (helicase) leads to destabilization of the Pol I – RNA complex , which induces termination of the elongation process (Braglia et al. 2010; Braglia et al. 2011; Kim et al. 2004; West 2004). Furthermore, the smallest subunit of Pol I (A12.2) is important for successful termination, as it putatively stimulates the intrinsic RNA cleavage activity of Pol I, that was previously mentioned. Aside from that, several other factors that share homology with RNA cleavage enhancing factor TFIIS have been shown to play a role in termination by Pol I (Prescott et al. 2004). After successful termination and release of the rDNA, Pol I is available for activation by Rrn3 and re-initiation of the next round of transcription.

2.3 Pol I pre-initiation complex formation

In *S. cerevisiae*, transcription initiation by RNA Polymerase I relies on four transcription factors. However, a basal initiation system of only three components has been shown to be sufficient for low level transcription of the rDNA gene in vitro *(Figure 3 B)*. For this basal system, in addition to Pol I only two transcription factors are required (Keener et al. 1997; Keener et al. 1998). Rrn3 has to bind to the stalk (subunit A14 / A43) of Pol I in order to render the polymerase active *(Figure 2)* (Blattner et al. 2011; Hori et al. 2023; Milkereit and Tschochner 1998; Neyer et al. 2016; Peyroche et al. 2000; Torreira et al. 2017)

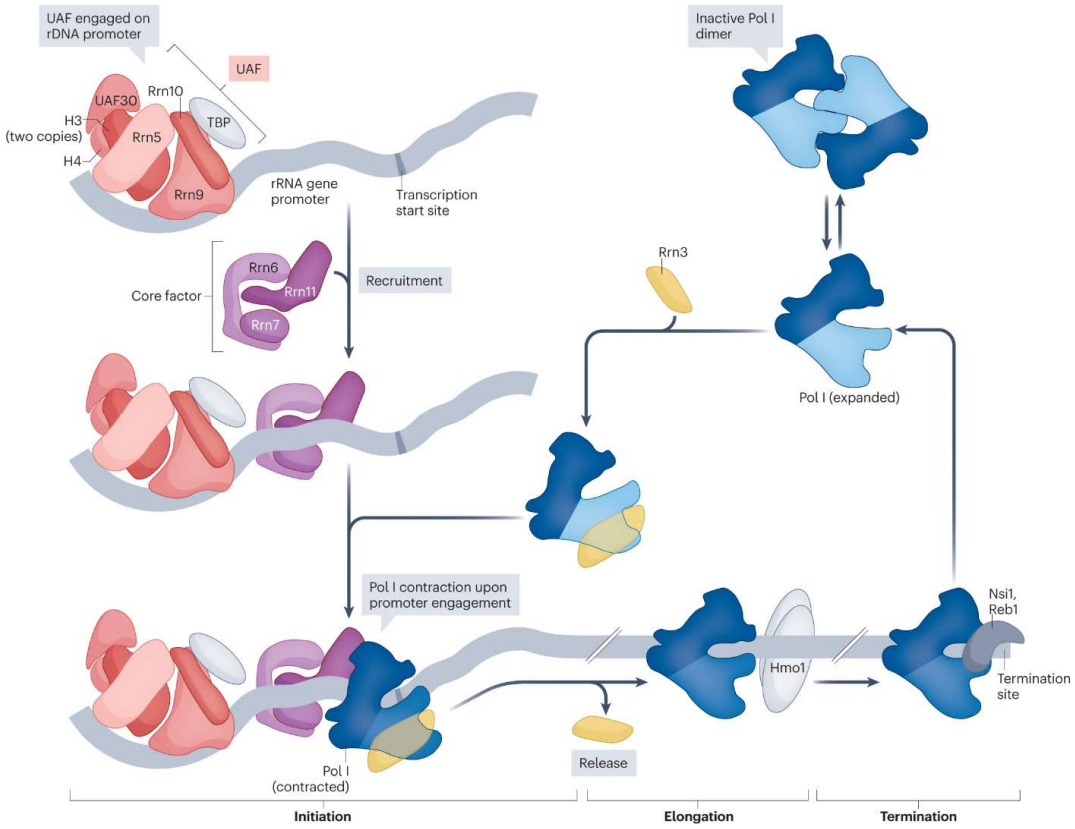


Figure 2: Schematic overview of RNA Pol I transcription initiation.

Simplified depiction of the transcription cycle with focus on transcription initiation .The rDNA promoter is dedicated to transcription by Pol I after binding of the six subunit Upstream Activating Factor – UAF. Subsequent assembly of TBP and Core Factor facilitates recruitment of monomeric, Rrn3-bound Pol I to the transcription start site. Pol I – DNA binding leads to contraction of the enzyme, which results in melting of the promoter DNA. The pre-initiation complex disassembles and Pol I engages elongation, which is assisted by high mobility group protein 1 – Hmo1. At the termination site, Pol I is released and can either hibernate in its dimeric form, or re-engage transcription after activation by Rrn3 binding (Hori et al. 2023).

This activation of Pol I by binding of Rrn3 is one of the key regulatory steps for rDNA transcription and is dependent on the phosphorylation state of both proteins. Although promoter binding can occur without Rrn3 attached, initiation can only take place when polymerase is activated (Bier et al. 2004; Cavanaugh et al. 2002; Claypool et al. 2004; Fath et al. 2001; Mayer et al. 2004). Recruitment of activated, Rrn3 – bound Pol I to the transcription start site (TSS) is facilitated by Core Factor (CF), a heterotrimeric complex, comprising Rrn6, Rrn7 and Rrn11 (Keys et al. 1994; Lalo et al. 1996), which binds to the Core Element (CE), stretching up to -38 nucleotides upstream of the TSS (Choe et al. 1992; Gubbey 2017; Pilsl et al. 2020). The DNA is tightly clamped between the promoter associated domain (PAD) of Rrn7 and the cyclin domains of Rrn11 on opposite sides of the core element, leading to a ~

30° bend in the promoter DNA (Pilsl et al. 2020). The exact role of this specific bend remains unclear. It is however assumed that bendability of the promoter DNA, rather than the sequence, plays a role for transcription initiation by enabling formation of three polymerase interacting regions (PIRs) between CF and Pol I (Engel et al. 2017; Engel et al. 2018; Gubbey 2017; Pilsl et al. 2020; Sadian et al. 2019).

Native levels of rDNA transcription require two additional, yet non - essential factors, the TATA Binding Protein (TBP) and the Upstream Activating Factor (UAF) (Keener et al. 1998; Steffan et al. 1996, 1998). First, UAF forms a stable complex with the upstream element (UE), spanning from -155 to -38, and commits the promoter to transcription initiation by Pol I. Following this, TBP binds to UAF *(Figure 2)*, which bridges a connection to the CF. Together, both proteins have a strong enhancing effect (up to 40 fold) on transcription in vitro and in vivo *(Figure 3 B)* (Keys et al. 1996; Steffan et al. 1996, 1998; Vu et al. 1999).

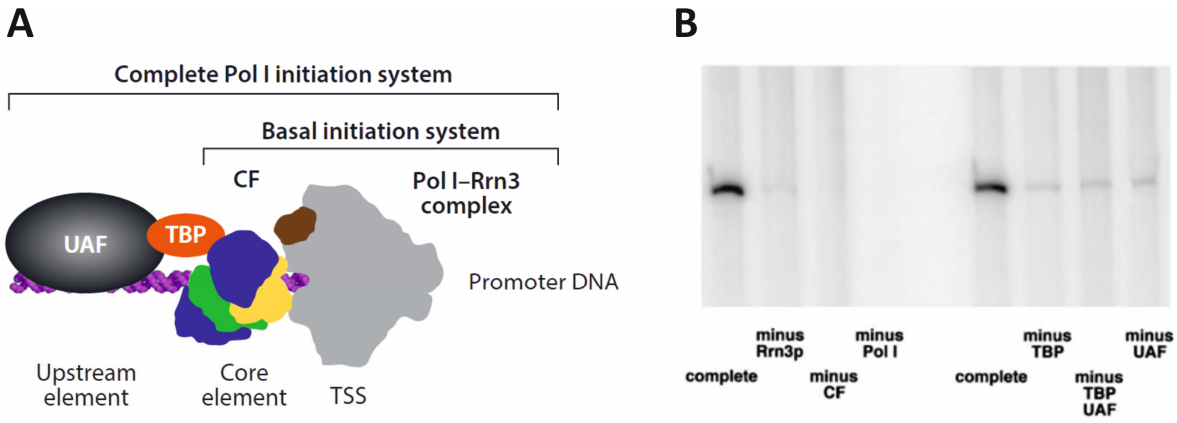


Figure 3: Comparison of the complete and basal RNA Pol I initation system.

A) Schematic representation of the complete initation system, comprising UAF, TBP, CF, Rrn3, and Pol I vs. the basal initation system, comprising only CF, Rrn3 and Pol I. (Engel et al. 2018) **B)** In vitro transcription assay done with purified components (UAF from yeast) showing the effect of omission of the different transcription factors. The basal initiation system represents the minimal requirements for initation. Exclusion of Rrn3, Core Factor (CF), or Pol I leads to a total loss of transcription (left half). In contrast, TBP and UAF are non-essential and transcription can take place at a very low rate (right half) (Keener et al. 1998).

Although it has been reported that overexpression of TBP can enhance transcription rates in absence of UAF (Aprikian et al. 2000), other publications refuted this hypothesis and demonstrate, that an overexpression of TBP cannot counteract loss of UAF in vivo or in vitro (Bedwell et al. 2012; Keener et al. 1998). In fact, they show that, in strains that lack UAF and

which overexpress TBP, a so-called polymerase switch (PSW) occurs, with leads to rDNA transcription relying fully on Pol II (Siddiqi et al. 2001). This phenomenon will be discussed in more detail later. Vica versa, a lack of TBP leads to transcription rates comparable to the level of the basal initiation system (Keener et al. 1998). This data indicates that healthy transcription levels can only be achieved with the assembly of the whole preinitiation complex, including all four factors in addition to Pol I *(Figure 3 B)*.

In early 2022, a cryo EM reconstruction of UAF and TBP in complex with rDNA promoter DNA was published (Baudin et al. 2022). This publication answered some long-standing questions about this crucial part of the initiation complex. This model and earlier studies which explored the interplay between UAF and TBP, suggest a bridging function of TBP, connecting the other two factors - CF and UAF (Engel et al. 2018; Keener et al. 1998).

As explained, in the upstream direction, TBP interacts with UAF, which is assembled to the UE. The atomic model shows an interaction between subunits Rrn9 and Rrn10 with both the N and C terminal lobe of TBP *(Figure 4)*(Baudin et al. 2022). Generally, this goes in line with findings by several research groups, linking an interaction between Rrn9 and TBP to the transcriptional activity of the Pol I system (Cormack and Struhl 1992; Schultz, Reeder, and Hahn 1992; J. S. Steffan et al. 1996, 1998). Interestingly, in the current model, the DNA binding surface of TBP (TATA binding saddle) is sequestered by a hydrophobic interaction with the N terminal helix of Rrn9. This would prevent TBP from scanning for binding sites in the rDNA promoter. It is speculated that this could lead to a low probability of Pol II and III PIC formation, as this particular DNA binding domain also plays a role in these assemblies (Baudin et al. 2022) (Cormack et al. 1992; Vannini et al. 2012). This hypothesis would go well in line with the observed polymerase switch (PSW) in cases of dysfunctional or absent UAF, leading to rDNA transcription by Pol II, as TBP would be freely available for assembly of this alternative initiation system. The function of UAF as a silencer of Pol II transcription, preventing this polymerase switch (PSW), will be discussed in more detail in chapter *2.4.3*.

Structurally, the downstream interaction between TBP and CF is rather poorly understood. Attempts to gather Cryo EM data of the whole PIC resulted in only a high-resolution density of Pol I – Rrn3 and Core Factor (Pilsl et al. 2020), and it is also not part of the previously mentioned Cryo-EM reconstruction of UAF.

Functional studies, however, showed that TBP can interact with all three subunits of CF. The strongest interaction was shown to occur with Rrn6 (Steffan et al. 1996, 1998). X – link data

from our own lab also suggest a vicinity of TBP to every subunit of CF, with Rrn6 being the only subunit that also crosslinks to UAF (Rrn10). It should however be noted, that these crosslinking experiments were done in absence of Pol I (Pilsl 2021). It is speculated, that the interaction could also rely on an insertion of TBP into the cyclin domains of Rrn7. This interface could have a similar effect as the previously mentioned occupation of the DNA binding domain of TBP by subunit Rrn9 of UAF (Baudin et al. 2022; Engel et al. 2017). Again, this model would suggest that TBP bridges the interaction between UAF and the basal initiation system, rather than functioning as an essential, DNA binding, transcription factor, as it does in the Pol II and III initiation systems (Cormack et al. 1992; Vannini et al. 2012). In these other initiation systems, TBP contacts DNA via its TATA binding saddle to allow PIC formation. Interestingly, DNase footprinting studies have mapped a TBP binding site located inside the upstream element, possibly providing a basis for Pol II PIC formation. However, the affinity of TBP towards this region is markedly lower than the affinity of UAF towards the same region. Moreover, the footprinting data also shows the same protection pattern of DNA when comparing TBP to TBP in complex with UAF (Baudin et al. 2022). Another study used photo crosslinking in to investigate binding behavior of Pol I PIC components in *A. castellanii*. While TBP crosslinks were observed in the region from position -78 to +10, this interaction was not mediated by TBPs TATA binding saddle (Bric et al. 2004).

Given the ambiguity of the existing data, it remains unclear if the interaction of UAF and TBP in the atomic model represents the native mode of binding, also because, unlike in the Pol II and III system, TBP is sequestered by UAF instead of promoter DNA (Baudin et al. 2022; Ravarani et al. 2020). Yet, although there is still much room for exploration of this highly important interaction, this model represents the current state of structural knowledge about this interaction.

Another key player in the regulation of Pol I rDNA transcription in yeast is the nucleolar protein Net1. Together with the phosphatase Cdc14 and Sir2, which will later be explored in more detail, it forms the REgulator of Nucleolar silencing and Telophase - complex (RENT – complex) (Shou et al. 1999; Straight et al. 1999), which is reported to influence chromatin structure, likely to silence rDNA transcription by Pol II (Goetze et al. 2010). Furthermore, Net1 engages in intricate interactions with various regulatory factors and kinases, e.g. CK2 and TORC1, which together regulate the cytoplasmic pH (Devare et al. 2020), adding layers of complexity to the regulatory role of the RENT complex and Net1. Concerning this project,

the most important role of Net1 is its direct link to stimulation of Pol I transcription (Hannig et al. 2019; Pilsl 2021; Shou et al. 2001). While Net1 is not part of the basal or complete in vitro transcription initiation system that were previously introduced *(Figure 3),* recent studies have shown that its C terminal region of Net1 is required to maintain healthy cell growth and is also reported to be sufficient for Pol I transcription stimulation in vitro (Hannig et al. 2019). Its stimulating effect as an additional factor in in vitro transcription assays with purified components of the Pol I PIC has been extensively studied by Michael Pilsl from our own lab. Further, crosslinking experiments revealed vicinity of Net1-C to Rrn9, H3, Rrn10, Rrn5, and TBP, suggesting that Net1-C directly interacts with UAF and TBP in context of the complete initation system, supporting their role as enhancers of rDNA transcription (Pilsl 2021).

Interestingly, the acidic tail of the human Pol I transcription factor UBF (upstream binding factor) shows amino acid sequence similarities with Net1, leading to speculations about conserved functional properties of this structural element across species (Hannig et al. 2019). UBF, despite its similar sounding name, shares no structural similarities with the yeast factor UAF. However, even though the role of UBF is still under investigation, it is argued that it might partly fulfill functions that are covered by UAF and Hmo1 in the yeast system, as it also stimulates transcription rates and it was initially identified as a protein that, similar to UAF, interacts with the upstream element in metazoans (Bell et al. 1988;). However, unlike UAF, UBF is not enough to commit the promoter to transcription by Pol I. In higher eucaryotes, this function is covered by SL1, which has been shown to nucleate PIC formation. However, this is stimulated by presence of UBF and (Friedrich et al. 2005; Leblanc et al. 1993; O'Sullivan et al. 2002; Panov et al. 2006; Schnapp and Grummt 1991; Smith et al. 1990). Other studies then revealed, that UBF, unlike UAF, also binds to the core element and inside the transcribed region of the rDNA gene, regulating the number of active rDNA genes (Bell et al. 1988; Leblanc, Read, and Moss 1993; Mais et al. 2005; Pikaard et al. 1990; Sanij et al. 2008).

Several elements of the basal *S. cerevisiae* Pol I PIC (Pol I + Rrn3 + CF) are however conserved. TifIA, the human homolog of Rrn3, has been shown to interact with TafIB, which is the functional homolog of Rrn7 (Miller et al. 2001) and part of the human transcription factor SL1. Aside from TafIB, SL1 contains additional subunits, TafIA (Rrn11), TafIC (Rrn6), and TBP, covering yeast factors CF and TBP. Further, it includes the TBP associated factor 1D (Taf1D) and Taf12 (Denissov et al. 2007; Gorski et al. 2007; Russell et al. 2006).

Likewise, the transcription machinery of the three RNA polymerases of yeast share common features. This reflects in the composition of the polymerase itself, which share 5 core subunits and 5 common subunits, to form a similar catalytic core structure. It further reflects in the different factors that are associated with the different steps; initiation, elongation, and termination. Interestingly, many of the factors that are subunits of Pol I and III, have homologs that act as separated transcription factors in the Pol II system. The similarities and differences between the three systems have been extensively reviewed, and will not be explored in detail in this thesis (Engel, Neyer, and Cramer 2018; Khatter, Vorländer, and Müller 2017; Vannini and Cramer 2012)(Cramer et al. 2008; Engel et al. 2018; Khatter et al. 2017; Vannini and Cramer 2012).

2.4 The Upstream Activating Factor UAF

2.4.1 General structural information

Although many research groups put efforts into understanding structure and function of UAF, structural information was not available for the most part of the last decades. As mentioned above, in 2022, the publication of a cryo EM reconstruction of UAF in complex with TBP and promoter DNA eventually provided detailled insight into the atomic structure and the architecture of the complex *(Figure 4 A)* (Baudin et al. 2022).

However, characterization of the complex began much earlier. In the 90ies, studies revealed that UAF comprises six different subunits, the two histone proteins H3 and H4, and four additional subunits Rrn5, Rrn9, Rrn10, and Uaf30 (Keener et al. 1997; Keys et al. 1996; Siddiqi et al. 2001). The possibility of the presence of a second H3 molecule was first described in 2018 via native mass spectrometry, showing a divergence of actual molecular weight of the complex when compared to the expected molecular weight for a 1:1 stoichiometry of all subunits (Smith et al. 2018). Interestingly, in addition to the two histone proteins H3 and H4, domain predictions of the different UAF subunits mapped a histone like domain to the C terminal region of Rrn5. This led to speculations about a nucleosome like core structure, involving H3, H4, and at least one other subunit, Rrn5 (Keener et al. 1997; Smith et al. 2018). In 2022, the publication of the structure of UAF eventually confirmed the presence of two H3 molecules as part of a hexameric, nucleosome like core structure. Aside from the histone fold domain of Rrn5, two additional histone folds were newly mapped to subunits Rrn9 and

Rrn10 and identified as part of this UAF core complex. Comparison of the interacting histone folds of Rrn9 and Rrn10 revealed similarities to the H2B and H2A dimer of the yeast nucleosome and resemblance to the human TFIID and the yeast SAGA complex. Furthermore, the histone fold interface of Rrn5 and the proximal (DNA contacting) H3, resembles a H4 – H3 dimer. Similar to the structure formed by Rrn9 and Rrn10, this element also shows high similarity to parts of TFIID and the SAGA complex *(Figure 4 D)*. Overall, the hexameric core of UAF exhibits strong parallels to structures that are involved in RNA transcription by Pol II and Pol III (Baudin et al. 2022; Butryn et al. 2015; Papai et al. 2020), although, in contrast to TFIID and SAGA, UAF is not essential for initiation. As the only subunit of UAF that is not involved in the formation of the hexameric core complex, Uaf30 is located in the periphery, with its N terminal winged helix domain contacting Rrn5 *(Figure 4 A)*. Due to its high flexibility, the center and C terminal region of this subunit are not resolved in the present atomic model of the complex. However, several studies have suggested that Uaf30 (Hontz et al. 2008; Siddiqi et al. 2001), especially its C terminus (Iida and Kobayashi 2019) plays a vital role in promoter recognition and rDNA repeat maintenance. Both functions will be discussed in a later section.

Drop out experiments showed that none of the six UAF subunits seem to be essential for (partial) complex formation. Surprisingly, despite recent studies suggesting a vital role of the two histone proteins for assembly of the nucleosome like core in the center of UAF (Baudin et al. 2022), even deletion of the two histone proteins H3 and/or H4 does not result in a total failure of complex formation, as the three Rrn subunits evidently still seem to assemble and co-purify. Likewise, drop-outs of Rrn5, Rrn9, and Rrn10 resulted in co-purification of the remaining subunits respectively (Smith et al. 2018). However, severe growth defects where observed in strains that carried a deletion of individual UAF subunits, demonstrating the importance of presence of every single subunit for the functionality of UAF and cell fitness (Hontz et al. 2008; Oakes et al. 1999; I. Siddiqi et al. 2001; I. N. Siddiqi et al. 2001; Vu et al. 1999). The implications behind these phenotypes will be discussed in chapter *2.4.3*.

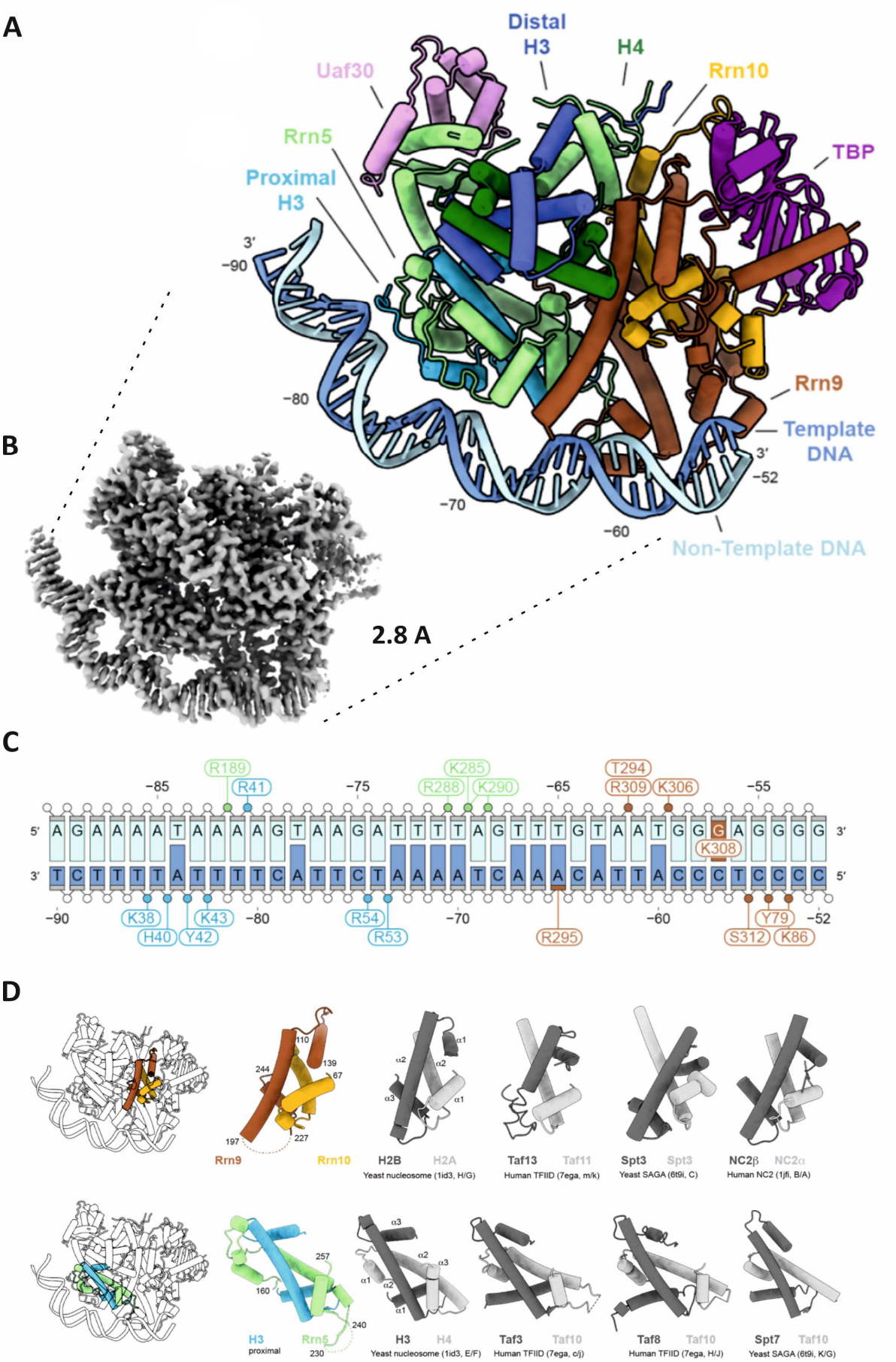

A
Distal H3
H4
Uaf30
Rrn10
Rrn5
TBP
Proximal H3
3'
-90
-80
Rrn9
Template DNA
3'
-52
-70
-60
Non-Template DNA
B
2.8 A
C
R189
R41
K285
R288
K290
T294
R309
K306
-85
-75
-65
-55
5' A G A A A A T A A A A G T A A G A T T T T A G T T T G T A A T G G G A G G G G 3'
3' T C T T T T A T T T T C A T T C T A A A A T C A A A C A T T A C C C T C C C C 5'
K308
-90
K38
H40
K43
Y42
-80
R54
R53
-70
R295
-60
Y79
S312
K86
-52
D
Rrn9
Rrn10
H2B
H2A
Yeast nucleosome (1id3, H/G)
Taf13
Taf11
Human TFIID (7ega, m/k)
Spt3
Spt3
Yeast SAGA (6t9i, C)
NC2β
NC2α
Human NC2 (1jfi, B/A)
H3
proximal
Rrn5
H3
H4
Yeast nucleosome (1id3, E/F)
Taf3
Taf10
Human TFIID (7ega, c/j)
Taf8
Taf10
Human TFIID (7ega, H/J)
Spt7
Taf10
Yeast SAGA (6t9i, K/G)

Figure 4: Structure of UAF and TBP bound to rDNA promoter - overview.

A) Simplified atomic model of the rDNA promoter (upstream element) bound by UAF and TBP. Helices are represented by cylinders, beta sheets by flat arrows. Together with two H3 and one H4 molecule, the histone folds of Rrn5, Rrn9, and Rrn10 form a hexameric core structure, while Uaf30 lies in the periphery of the core complex. **B)** Electron density @ 2.8 Å resolution of the model represented in A. **C)** Protein – DNA contacts (derived from the model) between the upstream element and UAF subunits Rrn5, H3 (proximal) and Rrn9. Colored dots represent unspecific phosphate backbone contacts, small colored rectangles show sugar contacts, colored bases depict base specific contacts. Contacts are colored according to colors in in A and C. **D)** Side by side comparison of histone-fold cores from the *S. cerevisiae* Pol I system (Rrn9/Rrn10 and H3/Rrn5) and yeast nucleosomes, yeast SAGA, and the human TFIID complexes. - (Baudin et al. 2022)

2.4.2 Promoterbinding

The current model reports a UAF - DNA interface that relies on interactions between the proximal (as opposed to the distal) H3 and Rrn5 in the middle region of the upstream element (UE) (position -86 / -68). In this region, specifically at position -78/-77, the DNA is strongly bent at a T-A dinucleotide, facilitating enhanced contacts with the minor groove in the direct viscinity, as also reported for DNA structure around nucleosome cores (Richmond et al. 1986). This interaction is supported by DNA contacts (mainly) between the C terminal part of Rrn9 and the downstream region of the UE (position -65 / -54). In this model the majority of all contacts are established with the phosphate backbone, with only one sugar and one base specific interaction by Rrn9 *(Figure 4 C)*. Excluding possible interactions between the unresolved fraction of Uaf30 and the promoter, the reported interfaces suggest a mostly sequence unspecific binding that is suggested to mainly rely on shape recognition and particular sequence elements that enable bendability (Baudin et al. 2022).

DNA binding and promoter targeting by UAF, however, was subject to various studies over the last decades, and a highly sequence specific interaction of UAF towards the rDNA promoter was already suggested in the 90ies (Keys et al. 1996). ChIP experiments showed that UAF fails to maintain its high specificity towards the upstream element when Uaf30 is omitted from the complex, which led to a decrease in template activation (Hontz et al. 2008). As a consequence, yeast cells will switch rDNA transcription systems (PSW strains), which will be discussed in the next chapter. This led to the conclusion that Uaf30 most likely plays a key role in efficient rDNA promoter recognition and subsequent transcription of the rDNA genes by Pol I. Further, comparison of UAF and a Uaf30 deletion complex in footprinting assays revealed loss of protection at an upstream promoter region roughly -113/-110 and -

107/-97 relative to the transcription start site (Hontz et al. 2008). Due to the limitations of these experiments, it was not clear if this effect was owed to structural changes in the complex after deletion of Uaf30, or lack of an actual interaction between Uaf30 and the UE. Other studies showed, that transcription rates were be susceptible to mutations in the C terminal region of Uaf30 (Iida and Kobayashi 2019). Interestingly, although this DNA region as well as most of Uaf30 was not resolved in the atomic model from 2022, electron density was detected between the promoter DNA template (position -100 to -92) and the N terminal domain of Uaf30 in a subpopulation of particles (Baudin et al. 2022) *(Figure 25)*. Taken together, the data suggests that Uaf30 plays an important role in sequence specificity and promoter targeting by UAF.

Aside from analysis of UAF or mutant complexes, the influence of the promoter sequence on UAF binding was extensively examined by different groups, including our own. While the data is not unanimous, in essence, several regions seem to play a role in UAF recruitment and specificity of the complex. Particularly the region around position -100 and further upstream relative to the transcription start site seems to partake in a specific interaction (Keener et al. 1998; Keys et al. 1996; Pilsl 2021) and stimulation of Pol I transcription (Kulkens et al. 1991; Musters et al. 1989).

2.4.3 UAF silences Pol II transcription

Although it is naturally part of the Pol I initiation system, which is responsible for rDNA transcription, it has been shown that **Pol I** is not absolutely necessary for this process. A so-called polymerase switch (PSW) can occur in mutant strains carrying dysfunctional UAF or Pol I (Conrad-Webb and Butow 1995; Hontz et al. 2008; Nogi et al. 1991; Oakes et al. 1999; Siddiqi et al. 2001; Siddiqi et al. 2001; Vu et al. 1999). While, as explored before, complex -, or rather subcomplex - formation seems to occur independently of the presence of every single subunit, functionality of UAF is strongly restricted if any of its subunits are missing. Sir2, as part of the RENT complex *(see 2.3)*, influences chromatin structure in a way that represses PIC formation and rDNA transcription by Pol II. Further, lack of functional or intact UAF has been shown to lead to loss of Sir2 from the rDNA. As a result, Pol II transcription is possible (Goetze et al. 2010). Although the Pol II promoter overlaps the Pol I promoter in the NTS2,

the transcription start site is identical, resulting in transcription of same rRNA precursor (Conrad-Webb and Butow 1995).

Interestingly, yeast strains lacking Uaf30, result in a slow growing phenotype that is able to transcribe rDNA using both Pol I and II. Although most of the transcription is still conducted by Pol I, Pol II transcription can range from ~ 5 - 14 %. Albeit at lower efficiency, the Pol I initiation system can still operate and initiate transcription under these circumstances. Silencing of Pol II PIC formation, however, seems slightly impaired (Hontz et al. 2008).

The phenotype of the Uaf30 deficient strain strongly differs from yeast strains with deletions of subunits Rrn5, Rrn9, or Rrn10, which suffer from an almost fatal growth defect that barely stabilizes without a helper plasmid . Interestingly, these deletion strains can bring forth clones that fully depend on a PSW aswell when cultivated over a longer time period (Oakes et al. 1999; Siddiqi et al. 2001; Vu et al. 1999).

The exact mechanism of Pol II silencing is not yet fully understood. However, a possible mechanisms, the inability to occupy TBP and restrict its ability to start Pol II PIC formation, was discussed before *(2.3)*. In light of the recently published structure of UAF, this could also explain why strains that carry a Uaf30 deletion are not as strongly affected as strains that lack one of the Rrn subunits, as partial complex formation without Uaf30 would still bring forth a complex that is able to occupy TBP via Rrn9/10 interactions that were introduced earlier *(2.3)*. Consequentially, disruption of the hexameric core complex by deletion of Rrn5, 9, or 10, could result in a damaged core structure that can no longer sequester TBP via Rrn9 and 10.

Aside from a dysfunctional UAF complex, the second prerequisite for the survival of PSW strains, especially for the ΔRrn – or Rrn defective strains, is an expansion of rDNA tandem repeat number. Interestingly, despite the reduction of rRNA output by ~70 % in Uaf30 defective mutants, the number of rDNA repeats is increased to roughly 200 repeats, and actively transcribed genes are highly occupied with polymerases (~ +100 % increase; Pol I and II) (Hontz et al. 2008; Siddiqi et al. 2001). It is however unclear, if Pol I and II transcripts derive from strains in which both polymerases can assemble to the same gene, different genes, or if heterogenous cultures are the reason for this. It is discussed, that the decrease in specificity due to loss of Uaf30 possibly leads to less promoters that are dedicated to Pol I transcription, leaving others open for Pol II PIC formation.

In contrast to these ΔUaf30 strains, which are still able to utilize Pol I for rDNA transcription, the Rrn deletion strains are highly dependent on a strong locus expansion in order to survive. In Rrn5 deficient strains, for example, repeat numbers of up to 400 – 700 copies were observed (Sasano et al. 2017).

The upregulation of rDNA copy numbers has its roots in the secondary target of UAF in the yeast genome, the SIR2 promoter. The function of UAF as a regulator of the rDNA repeat number will be discussed in the next chapter.

2.4.4 UAF regulates rDNA repeat count

In yeast, under healthy conditions, the rDNA locus spans over roughly 150 repeats of the rDNA gene, of which only around 50 % are actively transcribed at a time (Dammann et al. 1993; French et al. 2003; Schweizer et al. 1969). The repetitive architecture of this locus increases the likelihood of a decrease in copy number through recombination between complementary regions (Kobayashi et al. 2004). UAF was identified as part of a counteracting regulatory loop including at least two additional proteins, Sir2 and Fob1, which function as an inhibitor or enhancer of locus expansion respectively. A simplified model for this regulatory process is depicted in *(Figure 5)*.

Identification of UAF as a protagonist in this rDNA repeat maintenance mechanism was a result of random mutagenesis and growth analysis of the randomly mutated low copy (rDNA repeats) yeast strains. Strikingly, the most detrimental mutations could all be traced back to the genes of three UAF subunits, Rrn5, Rrn9 and Rrn10 (NGS). Furthermore, mutations in the Uaf30 gene resulted in a growth defect, but were less impactful than mutations in the Rrn genes. However, the research group also reported that Sir2 repression negatively correlates with mutations particularly in the C terminal region of Uaf30, implying a special role of the C terminal domain of Uaf30 in the regulation of Sir2 expression and thus in maintaining healthy rDNA copy numbers (Iida and Kobayashi 2019).

Fob1 (DNA replication fork – blocking protein) targets E-pros that are located in the IGS of the rDNA genes. E-pros are unidirectional noncoding Pol II dependent promoters, that control a cis-acting factor called EXP (E – pros: EXP promoters). When E-pros are activated during replication of NTS1 and NTS2 *(Figure 1 B)* Fob1 binds to these promoters and blocks passage of the replication fork at the replication fork barrier (RFB), preventing the fork from

clashing with the 35 S rDNA transcribing Pol I. This, however, can lead to doublestrand breaks (Brewer et al. 1992; Kobayashi et al. 2004; Takeuchi et al. 2003). These double strand breaks are more prone to recombination with unequal sister chromatids, because of reduced cohesion association in these regions. Hence, the repairing process of doublestrand breaks in this region is the main driver for the upscaling of rDNA repeats (Burkhalter and Sogo 2004; Hori et al. 2023; Iida and Kobayashi 2019; Kobayashi 2011; Kobayashi et al. 2004; Kobayashi and Ganley 2005).

Fob1 binding is usually repressed by the NAD – dependent histone deacetylase Sir2, which downregulates E-pro activation. This repression leads to a stable number of repeats by obstructing doublestrand breaks by Fob1 and, by this, prevents recombination and locus expansion. In line with this, deletion of the FOB1 gene has been shown to render rDNA repeat number inalterable (Defossez et al. 1999; Johzuka and Horiuchi 2002; Kobayashi et al. 1998). The direct correlation between rDNA repeat number and Sir2 on protein and RNA levels, was shown via WB and q-RT PCR respectively (Iida and Kobayashi 2019). Sir2 expression, in turn, is downregulated by excess UAF, which has a cell cycle independent, consistently low prevalence in the cell (Kulak et al. 2014). UAF has a higher binding affinity towards its primary target, the upstream element. However, as a consequence of repeat loss, free UAF increasingly binds to the Sir2 promoter, repressing its transcription in a concentration dependent manner. The reduced prevalence of Sir2 eventually leads to more Fob1 induced double strand breaks, and by this, to amplification of the rDNA repeat number. Interestingly, ChIP experiments demonstrated, that the SIR2 promoter seems to be the only other UAF binding site in the *S. cerevisiae* genome apart from the rDNA promoter (Iida and Kobayashi 2019).

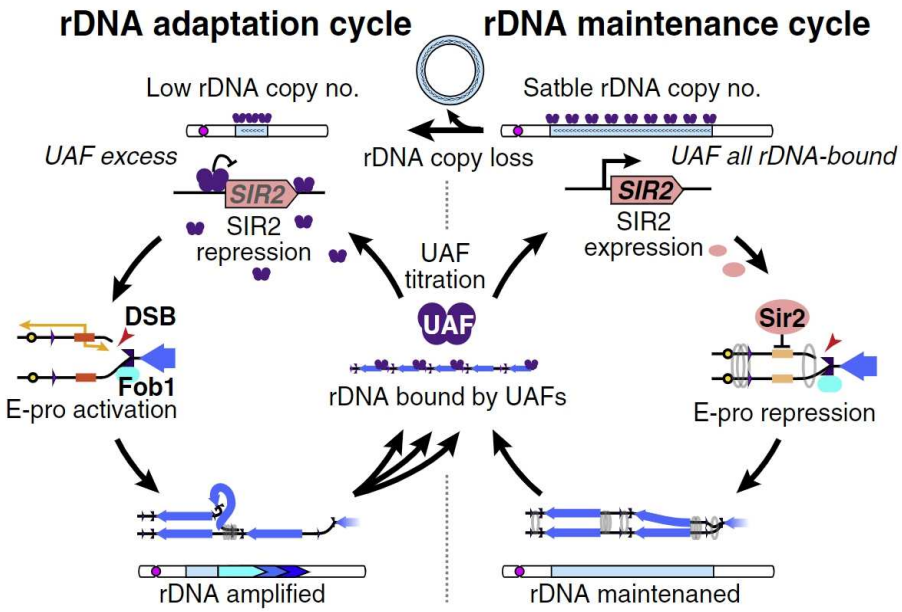


Figure 5: The role of UAF in the Sir2 rDNA repeat maintenance cycle.

UAF functions as a regulatory element in the upregulation of rDNA repeats via recombination between unequal sister chromatids. Loss of repeats leads to excess UAF which can bind to its secondary target, the Sir2 promoter, where it functions as a suppressor of SIR2 transcription. Binding of Fob1, which leads to doublestrand breaks (DSB), is no longer repressed by Sir2. Repairing of these DSBs leads to recombination and amplification of the rDNA repeat number (Iida and Kobayashi 2019)

2.5 Scope of this project

It is undisputed that the upstream activating factor 'UAF' plays a pivotal role as part of the rDNA transcription machinery of *Saccharomyces cerevisiae*, contributing to ribosomal homeostasis and, ultimately, cell fitness. Key functions as enhancer for Pol I transcription (Keener et al. 1997; D. A. Keys et al. 1996; Steffan et al. 1996), silencer of Pol II transcription (Oakes et al. 1999; Vu et al. 1999), and as a key regulator in the maintenance cycle of the rDNA repeat number (Hori et al. 2023; Iida and Kobayashi 2019; Kobayashi 2011) have been extensively studied. Structural data, however, was completely missing when this project started in 2018.

With this project, we aimed at providing complementary functional data, as well as structural data on UAF subunits. Employing various DNA binding assays like EMSA and fluorescence anisotropy, we wanted to adress questions regarding the DNA binding characteristics of individual UAF subunits, as well as domains. For this purpose, numerous constructs were generated and expression and purification protocols were optimized in a way that maximized yield and purity of each protein. Construct optimization led to additional generation of several truncated versions of the full lengths subunits, which were also characterized. Characterization was further extended to the whole UAF complex, as well as deletion mutants lacking two DNA binding interfaces, one of which was newly identified during this project. In addition to their DNA binding properties, the different UAF variants were subjected to testing for their ability to stimulate Pol I transcription initiation via in vitro transcription assays.

The structural characterization followed a bottom-up approach. With the crystallization of UAF subunits and domains, as well as the whole complex, we aimed at incrementally unveiling fragments of the whole complex. To achieve this, high throughput crystallization screenings were carried out for all suitable constructs.

3 Material & Methods

3.1 Materials

3.1.1 Strains

Table 1: Bacterial strains used in this study

Strains	Genotype	Supplier
E. Coli XL1blue	rec1A, endA1, gyrA96 thi-1, hsdR17 supE44, relA1, lac [F′proAB lacIqZΔM15 Tn10, TetR]	Stratagene
E. Coli BL21 CodonPlus (DE3) RIL	B F– ompT hsdS(rB – mB –) dcm+ Tetr gal λ(DE3) endA Hte [argU, ileY, leuW, CamR]	Stratagene

3.1.2 Oligonucleotides

Table 2: Oligonucleotides used in this study

PB#	Oligos	Sequence	Comment
Representation of all oligos used for constructs that are shown in the results of this thesis. QC-pair = quickchange pair (+ strand), V-ov. = Vectorspecific recombination overhang included, GS-ov. = gene specific recombination overhang included, Seq = sequencing primer, AF = annealing fragment			
03	Rrn9_N197Stop	GCATGATAAAATAGCCAAAGAGTAGGAATTCGATGTAAGGCAAG	qc-pair
04	Rrn9_P259Stop	GAGCTATATAACGATATCTGAGAAAAGTATAAAAAGAG	qc-pair
05	Rrn9_P280Stop	CTAAAAAAGTATCACCAATGAAAAAAGACCAGTTC	qc-pair
06	Rrn9_F335Stop	GCTTTAAACAAAAGAACATGATTTCAGGTAAAGGGC	qc-pair
07	Rrn10_Q120Stop	CCAGACAGAACAGGGAAGGTGAAGTAAGGTAATCGGGAAGGG	qc-pair
98	Rrn9_172_qc1	GATGAATTGAACATTCCATAGGAAATATCCCGGAATATC	qc-pair
99	Rrn9_172_qc2	GATATTCCGGGATATTTCCTATGGAATGTTCAATTCATC	
09	Chth_Uaf30_fw	AAGTTCTGTTTCAGGGCCCG ATGAGTACCCAGCTTACAGAAG	V-ov.
10	Chth_Uaf30_revc	ATGGTCTAGAAAGCTTTA CTCATCAATGGGGTACAACTGG	V-ov.
11	Chth_Rrn5_T221fw	AAGTTCTGTTTCAGGGCCCG ACATCATACAATCCTAGTTACTTAGCC	V-ov.
13	Chth_Rrn5_D315rev	ATGGTCTAGAAAGCTTTA ATCATCACGTTTCCATTTGTCACG	V-ov.
124	UAF_Co_B_1_fw	AAGTTCTGTTTCAGGGCCCG AGTGATCTTGACGAAGAAAGTCAAATTG	V-ov.
125	UAF_Co_B_1_rv	TCATATGTTCCCATTAGGCAGTTCTATG	GS-ov.
126	UAF_Co_B_2_fw	CTAATGGGAACATATGATCGAACAGAAAGTAATCGTATTGTACACG	GS-ov.
127	UAF_Co_B_2_rev	GTGATGGCTGCTGCCCATTATATCTCCTTCTTATACTTAACTAATATACTAAGATGGGG	GS-ov.
128	UAF_Co_B_3_fw	ATGGGCAGCAGCCATCAC	GS-ov.
129	UAF_Co_B_3_rv	TTAAATGCCCTTTGTATCTGATTGCTC	GS-ov.
130	UAF_Co_B_4_fw	GATACAAAGGGCATTTAA ACGATGCGTCCGGCGT	GS-ov.
131	UAF_Co_B_4_rev	ATGGTCTAGAAAGCTTTA GATATTACCTGGCGCATGCCTATAATC	V-ov.
132	UAF_Co_pCDF_1_fw	TAATAAGGAGATATACCATG GCCAGAACAAAGCAAACAGCAAG	V-ov.
133	UAF_Co_pCDF_1_rv	CTATGATCTTTCACCTCTTAATCTTCTAGCC	GS-ov.

134	UAF_Co_pCDF_2_fw	GGTGAAAGATCATAG ACGATGCGTCCGGCGT	GS-ov.
135	UAF_Co_pCDF_2_rv	TTATTTGCTCAGCCATTTCAGCAG	GS-ov.
136	UAF_Co_pCDF_3_fw	GGCTGAGCAAATAATCGAACAGAAAGTAATCGTATTGTACACG	GS-ov.
137	UAF_Co_pCDF_3_rv	CCTCTACCGGACATTATATCTCCTTCTTATACTTAACTAATATACTA AGATGGGG	GS-ov.
138	UAF_Co_pCDF_4_fw	ATGTCCGGTAGAGGTAAAGGTGG	GS-ov.
139	UAF_Co_pCDF_4_rv	TTCTTTACCAGACTCGAGTTA ACCACCGAAACCGTATAAGGTTC	V-ov.
140	UAF_Co_sq1	GATGTCGGCGATATAG	Seq.
141	UAF_Co_sq2	TCCAAGCTACAACGC	Seq.
142	UAF_Co_sq3	ACGACGACTCTGGGC	Seq.
143	UAF_Co_sq4	CGAGATCGATCTCGATC	Seq.
144	UAF_Co_sq5	CAAGAAAACCAAGAG	Seq.
148	UAF_Co_sq6	CGTGTAGTACACTAC	Seq.
149	UAF_Co_sq7	CAGCTGCTGATGTTC	Seq.
150	UAF_Co_sq8	CTGGACAGCTTATTC	Seq.
151	UAF_Co_sq9	GATCGAGATCTCGATC	Seq.
152	UAF_Co_sq10	CGATTATTTGCTCAGC	Seq.
153	Rrn9_dN29_Bfw	AAGTTCTGTTTCAGGGCCCGACAACCGCTGATAAAGAG	V-ov.
159	CE_+8rev_Cy3	TTTCGCATGAAGTACCTC	
111 112	Frag 1 + Frag 1 -	TTTTTCTCGGCGAGAAATACGTAGTTAAGGCAGAGCGACA TGTCGCTCTGCCTTAACTACGTATTTCTCGCCGAGAAAAA	AF
113 114	Frag 2 + Frag 2 -	CAGAGCGACAGAGAGGGCAAAAGAAAATAAAAGTAAGATT AATCTTACTTTTATTTTCTTTTGCCCTCTCTGTCGCTCTG	AF
115 116	Frag 3 + Frag 3 -	AAGTAAGATTTTAGTTTGTAATGGGAGGGGGGGTTTAGTC GACTAAACCCCCCCTCCCATTACAAACTAAAATCTTACTT	AF
117 118	Frag 4 + Frag 4 -	GGGTTTAGTCATGGAGTACAAGTGTGAGGAAAAGTAGTTG CAACTACTTTTCCTCACACTTGTACTCCATGACTAAACCC	AF
170	dUaf30UAF_fw	GGGAACATATGATCGCGTAGAAAGGAAAAGCCCATCGTTTC	qc-pair
171	dUaf30UAF_rev	CGATCATATGTTCCCATTAGGCAGTTC	qc-pair
104	-183 fw	CGCTAAGATTTTTGGAGAAT	
	UE fwd	AGCTTAAATTGAAGTTTTTCTC	MP
	PIP_Reb1_fwd	TTACCCGGGGCACCTGTC	MP 4228
	PIP_Reb1_Cy5	TTACCCGGGGCACCTGTC	MP 4229
	Cy3 PIP_rev_119 nt	ATCACCTAGCGACTCTCTCC	MP 4404
	PIP_rev_119 nt	ATCACCTAGCGACTCTCTCC	MP 4405
	147 nt rev	GCATTCTCGAGACGGTGTAG	MP 4407

3.1.3 Plasmids

Table 3: Bacterial plasmids used in this study

#	Insert	Vector	Origin
Representation of all bacterial vectors used in this thesis. Plasmids marked with origins CE and MP derive from previous projects by Christoph Engel and Michael Pilsl respectively. pOPIN B vectors: HRV-3C cleavage site after N terminal His tag. pCDF: non cleavable N term His tagMP pEX vectors were used for S. cerevisiae promoter template amplifications. P – numeration corresponds to own vector library, all other vectors correspond to the Tschochner lab library.			
P5	His – Rrn9 – fl / Rrn10 – fl	pOPIN B	CE
P6	His – Uaf30 – fl	pCDFduet-1	CE
P7	His – Uaf30 – Ntd	pOPIN B	CE
P8	His – Uaf30 – Ctd	pOPIN B	CE
P12	His – ct. Rrn5like – Ntd	pOPIN B	this project
P16	His – ct. Uaf30like – fl	pOPIN B	this project
P18	His - ΔC168 Rrn9 / ΔC25 Rrn10	pOPIN B	this project
P56	His - ΔC168 Rrn9 / Rrn10 – fl	pOPIN B	this project
P57	His - ΔC105 Rrn9 / Rrn10 – fl	pOPIN B	this project
P58	His - ΔC85 Rrn9 / Rrn10 – fl	pOPIN B	this project
P59	His - ΔC30 Rrn9 / Rrn10 – fl	pOPIN B	this project
P61	His – Rrn9 / His – Uaf30 / His – Rrn10	pOPIN B	this project
P62	H3 / His – Rrn5 / H3	pCDFduet-1	this project
P63	His - ΔN29 Rrn9 / Rrn10 – fl	pOPIN B	this project
P64	His – ΔC85 Rrn9 / His – Uaf30 / His – Rrn10	pOPIN B	this project
P65	His – Rrn9 / His – Rrn10	pOPIN B	this project
2517	PIP WT promoter	pEX	MP pEX A2 2
2527	PIP Δ UE (-155 -39)	pEX	MP pEX A2 12

3.1.4 Enzymes

Table 4: Enzymes used in this study

Enzymes	Usage	Supplier
NcoI	pOPIN B linearization	NEB
KpnI	pOPIN B linearization	NEB
HindIII	pOPIN E linearization	NEB
PmeI	pOPIN E linearization	NEB
DpnI	Digestion of met. DNA / mutag. PCR	NEB
HRV-3C protease	Tag cleavage	Lab-own
Phusion Polymerase	PCR	Lab-own
Proteinase K	Digest of proteins in Trx assays	Sigma
5x InFusion HD Enzyme mix	InFusion cloning	Takara
DNase	Bacterial cell lysis	Invitrogen

3.1.5 Buffers

Table 5: Buffers used in this study

Buffer	Composition
All purification buffers were used with a final 1 mM DTT. Lysis buffers were used with a final concentration of 1x protease inhibitor mix. Both were added prior to buffer usage.	
Uaf30 Lysis	350 mM NaCl; 20 mM Imidazole; 10% glycerol; 20 mM Hepes pH 7.8; 1 mM $MgCl_2$
Uaf30 W50	200 mM NaCl; 50 mM Imidazole; 10% glycerol; 20 mM Hepes pH 7.8; 1 mM $MgCl_2$
Uaf30 Highsalt	1 M NaCl, 50 mM Imidazole, 10 % glycerol, 20 mM HEPES pH 7.8; 1 mM $MgCl_2$, 1 mM DTT
Uaf30 W75	200 mM NaCl; 75 mM Imidazole; 10% glycerol; 20 mM Hepes pH 7.8; 1 mM $MgCl_2$
Uaf30 Elution	200 mM NaCl; 350 mM Imidazole; 10% glycerol; 20 mM Hepes pH 7.8; 1 mM $MgCl_2$
Uaf30 SEC	150 mM NaCl; 20 mM Hepes pH 7.8; 1 mM $MgCl_2$
Uaf30 N/C Lysis	350 mM NaCl; 20 mM Imidazole; 10% glycerol; 20 mM Hepes pH 7.8; 1 mM $MgCl_2$
Uaf30 N/C Wash	200 mM NaCl; 50 mM Imidazole; 10% glycerol; 20 mM Hepes pH 7.8; 1 mM $MgCl_2$
Uaf30 N/C Elution	200 mM NaCl; 350 mM Imidazole; 10% glycerol; 20 mM Hepes pH 7.8; 1 mM $MgCl_2$
Uaf30 N Hep 200	200 mM NaCl; 10% glycerol; 20 mM Hepes pH 7.8; 1 mM $MgCl_2$
Uaf30 N Hep 1000	1000 mM NaCl; 10% glycerol; 20 mM Hepes pH 7.8 1 mM $MgCl_2$
Uaf30 N/C SEC	150 mM NaCl; 20 mM Hepes pH 7.8; 1 mM $MgCl_2$
Rrn9/10 Lysis	500 mM NaCl; 20 mM Imidazole; 10% glycerol; 20 mM Hepes pH 7.8; 1 mM $MgCl_2$
Rrn9/10 W50	350 mM NaCl; 50 mM Imidazole; 10% glycerol; 20 mM Hepes pH 7.8; 1 mM $MgCl_2$
Rrn9/10 W75	350 mM NaCl; 75 mM Imidazole; 10% glycerol; 20 mM Hepes pH 7.8; 1 mM $MgCl_2$
Rrn9/10 Elution	350 mM NaCl; 350 mM Imidazole; 10% glycerol; 20 mM Hepes pH 7.8; 1 mM $MgCl_2$
Rrn9/10 SEC	300 mM NaCl; 20 mM Hepes pH 7.8; 1 mM $MgCl_2$
IEX200 binding buffer	200 mM NaCl, 20 mM HEPES pH 7.8, 10 % glycerol, 1 mM $MgCl_2$
UAF Lysis	400 mM $(NH_4)_2SO_4$; 200 mM Tris-HCl pH 8.0; 10 % glycerol
UAF Wash	450 mM KCl; 20 mM Imidazole; 20 mM Tris-HCl pH 8.0; 20 % glycerol
UAF ATP Wash	UAF Wash buffer + 5 mM ATP + denatured protein
UAF Elution	450 mM KCl; 20 mM Tris-HCl pH 8.0; 20 % glycerol
UAF S 350 (load)	350 mM KCl; 20 mM Tris-HCl pH 8.0; 20 % glycerol
UAF S 500 (wash)	500 mM KCl; 20 mM Tris-HCl pH 8.0; 20 % glycerol

UAF S 750 (elution)	750 mM KCl; 20 mM Tris-HCl pH 8.0; 20 % glycerol
UAF SEC	450 mM KCl; 20 mM Tris-HCl pH 8.0; 5% glycerol
C.t. Rrn5-like Lysis	350 mM NaCl; 20 mM Imidazole; 10% glycerol; 20 mM Hepes pH 7.8; 1 mM $MgCl_2$
C.t. Rrn5-like Wash	200 mM NaCl; 75 mM Imidazole; 10% glycerol; 20 mM Hepes pH 7.8; 1 mM $MgCl_2$
C.t. Rrn5-like Elution	200 mM NaCl; 350 mM Imidazole; 10% glycerol; 20 mM Hepes pH 7.8; 1 mM $MgCl_2$
C.t. Rrn5-like SEC	150 mM NaCl; 20 mM Hepes pH 7.8; 1 mM $MgCl_2$
5x EMSA / Anisotropy Reaction buffer	100 mM Tris-HCl pH 7.5; 0.5 mg/ml BSA; 5 mM DTT; 50 % (v/v) glycerol
OrangeG loading dye	60 % (v/v) glycerol; 0.4 % (w/v) OrangeG; 10 mM Tris-HCl pH 7.6; 60 mM EDTA
5x TRX (no NTPs)	20 mM Hepes-KOH pH 7.5; 10 mM $MgCl_2$; 5 mM EGTA; 0.05 mM EDTA; 2.5 mM DTT
5x TRX (NTPs)	20 mM Hepes-KOH pH 7.5; 10 mM $MgCl_2$; 5 mM EGTA; 0.05 mM EDTA; 2.5 mM DTT; 0.2 mM ATP / GTP / CTP; 0.05 mM CTP + 0.3 µCi α^{32}P-CTP (add prior to use / Hotlab)
TRX Proteinase K buffer	10 mM Tris-HCl pH 7.5; 300 mM NaCl; 0.55 % (w/v) SDS; 5 mM EDTA + 0,5 mg/ml Proteinase K (add prior to use) + 20 ng/µl glycogen (add prior to use)
TRX KAc dilution buffer	20 mM Hepes-KOH pH 7.5; 10 % (v/v) glycerol; 1 mM $MgCl_2$; 0.4 mg/ml BSA; 200 mM KAc; 5 mM DTT
TRX H0 buffer	10 mM Hepes-KOH pH 7.5; 10 % (v/v) glycerol; 2 mM $MgCl_2$; 0.1 mM EDTA; 2.5 mM DTT
RNA loading dye	0.1 x TBE, 80% (v/v) Formamide (deionized); 0.02 % (w/v) Bromphenolblue; 0.02 % (w/v) Xylencyanol
DNA annealing buffer	50 mM NaCl; 10 mM Tris-HCl pH 7.8
5x Protein loading dye (SDS)	130 mM Tris-HCl pH 6.8; 15 % (v/v) glycerol; 2.1 % (w/v) SDS; 0.15 % (w/v) Bromophenolblue; 5 % (v/v); β-Mercaptoethanol
4x Lower Tris (SDS PAGE)	1,5 Tris; 0.4 % (w/v) SDS; pH 8.8
4x Upper Tris (SDS PAGE)	0.5 M Tris; 0.4 % (w/v) SDS; pH 6.8
Separating gel	1x Lower Tris buffer; 20 % (v/v) Rotiphorese®Gel (30 % solution); 1 % (v/v) APS; 0.1 % (v/v) TEMED
Stacking gel	1x Upper Tris buffer; 40 % (v/v) Rotiphorese®Gel (30 % solution); 1 % (v/v) APS; 0.1 % (v/v) TEMED
SDS PAGE running buffer	50 mM MOPS; 50 mM Tris pH 7.7; 0.1 % (w/v) SDS; 1 mM EDTA
Native PAGE gel 10% / 6%	4 % (v/v) 10 x TBE
WB transfer buffer	25 mM Tris-HCl pH 8.5; 20 % (v/v) MeOH; 192 mM glycerol
WB blocking solution	2.5 % BSA / TBS-T 0.1%
WB antibody binding buffer	2 % BSA / TBS-T 0,1%
TBE (10x)	900 mM Tris; 900 mM Boric acid; 10 mM EDTA
TBS (1x)	20 mM Tris; 150 mM NaCl
TBS-T	TBS + 0.1 % (v/v) Tween
Tfb I	30 mM KAc; 50 mM $MnCl_2$; 100 mM KCl; 15 % glycerol

Tfb II	10 mM MOPS; 75 mM $CaCl_2$; 10 mM KCl; 15 % glycerol
UREA gel	1 x TBE; 23.3 % (v/v) Rotiphorese®Gel (30 % solution); 8 M UREA
Inoue Transformation Buffer	10 mM PIPES pH 6.7; 55 mM $MnCl_2$; 15 mM $CaCl_2$; 250 mM KCl
Phusion PCR Mastermix (Mona Höcherl)	0.16 mM dNTPs; Phusion Polymerase (lab own); 5x HF-buffer (NEB); Phu-Buffer 7

3.1.6 Kits

Table 6: Commercial kits used in this study

Kits	Supplier
Miniprep Kit	Thermo Fisher Scientific
Nucleospin® Gel and PCR Cleanup Kit	Macherey – Nagel
peqGOLD Plasmid Miniprep Kit	peqLab
Qiagen Spin Miniprep Kit	Qiagen
Qiagen PCR and Gel cleanup Kit	Qiagen

3.1.7 Chemicals

Table 7: Chemicals and chemical mixes used in this study

Chemicals	Supplier	Comment
If not stated otherwise, chemical stocks were solved in H_2O (Millipore).		
IPTG	Sigma Aldrich	
100x Protease Inhibitor Mix	-	Solved in 80 % EtOH
200 mM Benzamidine	Sigma Aldrich	
100 mM PMSF	Roth	
Imidazole	Sigma Aldrich	
SDS	Sigma Aldrich	
BM Chemiluminescence Blotting Substrate (POD)	Roche	
SYBR Safe DNA gel stain	Invitrogen / Thermo Fisher	Final conc.: 1 : 20,000
UREA	Sigma Aldrich	
Agarose	Sigma Aldrich	
DMSO	Life technologies	
Glycogen	Sigma Aldrich	
TEMED	Thermo Fisher Scientific	
APS	Roth	

3.1.8 Growth media

Table 8: Media used for bacterial cultivation

Medium/Plates	Composition	Comment
LB Agar	1 % (w/v) Tryptone 0.5 % (w/v); yeast extract; 1 % (w/v) NaCl; 2 % (w/v) Agar	
LB medium	1 % (w/v) Tryptone; 0.5 % (w/v) yeast extract; 1 % (w/v) NaCl	
TB medium	2.4 % (w/v) yeast extract; 2 % (w/v) Tryptone; 0.4 % (v/v) glycerol; H_2O ad 900 ml; 100 ml 10 x NPS (salt solution) and 50 ml 5052 autoinduction mix	Autoclave in 900 ml or 1800 ml volume; add sterile NPS and 5052 prior to use
SOB medium	10 mM NaCl; 2.5 mM KCl; 2 % (w/v) Tryptone; 0.5 % (w/v) Yeast Extract; 10 mM $MgCl_2$; 10 mM $MgSO_4$	
5052 autoind. mix	glycerol 25% (v/w); 0.14 M Glucose; 0.3 M α-Lactose	
20x NPS	1 M KH_2PO_4; 1 M K_2HPO_4; 0.5 M $(NH_4)_2PO_4$	

3.1.9 Antibiotics

Table 9: Antibiotics used for selection

Antibiotics	Final dilution / concentration
Ampicillin	1:1000 / 100 µg/ml
Chloramphenicol	1:1000 / 30 µg/ml
Kanamycin	1:1000 / 50 µg/ml
Streptomycin	1:1000 / 50 µg/ml

3.1.10 Consumables

Table 10: Consumables used in this study

Consumables	Manufacturer
Amicon centrifugal filters (MWCO 3 – 100 kDa)	Merck Millipore
PD-10 colums	GE Healthcare
IntelliPlate 96-3 lvr	Hampton Research
384 microplate, low volume, U bottom, black	Corning®
NuPAGE™ 10 – 12 % Bis-Tris gel	Invitrogen / Thermo Fisher
NuPAGE™ 4 – 12 Bis-Tris gradient gel	Invitrogen / Thermo Fisher
24 well pre greased crystallization plates	Crystalgen

3.1.11 Chromatography columns

Table 11: Chromatography columns used for protein purification

Columns	Manufacturer
HisTrap 1 / 5 ml	GE Healthcare / Cytiva
HiTrap Heparin 1 / 5 ml	GE Healthcare / Cytiva
HiTrap SP HP 1 ml	GE Healthcare / Cytiva
HiTrap Q HP 1 ml	GE Healthcare / Cytiva
Superdex 75 Increase 10/300 GL	GE Healthcare / Cytiva
Superdex 200 Increase 10/300 GL	GE Healthcare / Cytiva
Superose 6 Increase 10/300 GL	GE Healthcare / Cytiva

3.1.12 Protein Crystallization Screens

Table 12: Commercial high throughput screens used in this study

Screen	Manufacturer
Pre Crystallization Test (PCT)	Hampton Research
Ammonium Sulfate Grid Screen (48 cond.)	Hampton Research
Index HT-96	Hampton Research
JCSG Plus HT-96	Molecular Dimensions
Morpheus I	Molecular Dimensions
MPD Grid Screen (48 cond.)	Hampton Research
PACT premier™ HT-96	Molecular Dimensions
PEG Ion HT-96	Hampton Research
Wizard 1+2	Molecular Dimensions
Wizard 3+4	Molecular Dimensions

3.1.13 Miscellaneous

Table 13: Miscellaneous

Object	Application	Manufacturer / Supplier
Colored prestained marker (#77125)	SDS PAGE	NEB
Colored prestained marker (#77195)	SDS PAGE	NEB
InstantBlue ™	Coomassie staining	Expedeon / VWR
SYBR Safe DNA stain	Agarose / EMSAs	Thermo Fisher Scientific
α-^{32}P-CTP	Trx Assays	Hartmann Analytic
dNTPs	PCR	NEB
NTPs	Trx Assay	NEB
PVDF membrane	Western Blot	BioRad

BSA	various	Roth
α His – HRP (His-Probe)	Western Blot	Thermo Fisher Scientific

3.1.14 Software & online tools

Table 14: Software used for data aquiration and evaluation

Software / Online - Tool	Publisher
FIJI ImageJ	FIJI
GraphPad Prism 9.5.1	Dotmatics
CorelDraw Graphic Suite - 2020	Alludo
ApE – A plasmid Editor	
Chimera 1.14	RBVI
HHpred	MPI Toolkit
PsiPred	UCL Bioinformatics Group
Quick2D	MPI Toolkit
AlphaFold	European Bioinformatics Institute
Office (Powerpoint/Excel/Word)	Microsoft
SparkControl ™	TECAN Spark / FA
Mosquito ® software v.4.1	SPT Labtech

3.1.15 Hardware

Table 15: Hardware used in this study

Hardware	Manufacturer
Mosquito ® LCP nanodispenser	SPT Labtech
Typhoon ™ FLA-9500 imager	GE Healthcare
PCR cycler primus 25 advanced	PeqLab
ÄKTA pure 25	GE Healthcare
ÄKTA micro Ettan LC	GE Healthcare
Branson Sonifier 450	Emerson
LAS-3000 Imager	FujiFilm
NanoDrop ™ One	Thermo Fisher Scientific
Molecular Imager ® PharosFX ™ Plus	BioRad
Prometheus Panta	Nanotemper
TECAN Spark ® microplate reader	TECAN
TransBlot Turbo System	Bio-Rad
TwoMP Mass Photometer	Refeyn
Branson Sonifier 450	Emerson

3.2 Molecular biology methods

3.2.1 Plasmid purification

Plasmid purification was achieved by extraction from transformed *E. Coli* XL1 blue competent cells. For this purpose, 5 - 10 ml of LB were inoculated from an LB plate or directly from the transformation medium, antibiotics were added and the cultures were grown at 37°C shaking for at least 14 h. For purification, all steps were done according to the Qiagen or PeqLab plasmid purification kit protocol.

3.2.2 Restriction digest

Plasmids were digested using restriction enzymes from New England Biolabs. Enzyme units were adjusted to the amount of DNA to be digested according to the respective NEB suggestions. Digestion was carried out at 37°C for at least 1 h. Analytical digests were checked via agarose gel electrophoresis. Preparative digests were subsequently gel purified following the Qiagen or PeqLab gel purification kit protocol.

3.2.3 PCR

Amplification of DNA via PCR was generally performed using a lab stock of Phusion Polymerase mastermix containing dNTPs as well as lab own Phusion polymerase with. 0.5 µM of each primer and 50 ng of the respective DNA template were added to each reaction tube. Reactions were filled up with H2O to a final volume of 50 µl. Annealing temperature and elongation time were adapted to fit the need of each particular amplification. Elongation times were set to 30 s per 1000 kb, annealing temperatures were set at least 5°C below the primer melting temperatures.

3.2.4 Purification of nucleic acids

Purification of PCR products or annealed DNA Fragments was achieved via PCR purification or gel extraction depending on the application of the respective DNA. For this, the instruction manuals of commercially available kits from Qiagen or PeqLab was followed. Elution was done in H2O Millipore.

3.2.5 In – Fusion Cloning

Introduction of inserts into the desired vectors was achieved via recombination using the In-Fusion enzyme mix (Takara). Forward and reverse primers of each insert were designed to have a 15 nt long overhang complementary to the 3' or 5' ends of the linearized target vector. In-Fusion reactions contained 50 ng of the linearized vector and 25 – 50 ng of insert, with adaptions depending on their length ratio. 1 µl of the In-Fusion enzyme mix was added and each reaction was supplemented with H_2O up to a final volume of 10 µl. Reactions were then incubated for 10 min at 50°C. For subsequent bacterial transformation, 2.5 µl of the In-Fusion reaction were added to 50 µl of chemically competent XL1blue cells.

In cases where more than one insert needed to be inserted, PCRs were performed to connect two fragments at a time by PCR amplification, using the respective forward and reverse primer of each terminal fragment. The connected fragments were finally inserted into their destination vector as stated above.

3.2.6 Site directed mutagenesis / quickchange PCR

Introduction of single or double nucleotide point mutations into vectors was achieved via quick-change PCR. Complementary primers carrying the desired mutation(s) roughly in the center of a ~30 bp stretch were therefore designed. PCR was performed as described in *3.2.3.* The annealing duration was increased to match the size of the respective plasmids. The product was subsequently digested with DpnI for 3 - 5 h at 37°C to remove original methylated DNA that was extracted from bacteria and used as an initial template for mutagenesis. Afterwards, the mutated vector was transformed into competent XL1 blue

cells. Clones were picked for Minipreps, and extracted plasmids were checked for the desired mutations via sanger sequencing (Microsynth Seqlab).

3.2.7 Agarose gelelectrophorese

0.8 – 2% agarose gels containing a final concentration of 1 : 20,000 (v/v) SYBR Safe were used to separate DNA. Samples were mixed with 6x DNA loading dye before application onto the gel. Gels were run at 120 V for 45 – 60 min depending on the length of the DNA.

3.2.8 Sequencing

Sequencing reactions were prepared to meet the requirements of Microsynth SeqLab. A minimum amount of 600 ng per plasmid were therefore premixed with the respective sequencing primers (2 µM final concentration) and H_2O Millipore to a final volume of 15 µl. Standard primers like T7 forward and reverse were added by Microsynth after receiving the samples.

3.2.9 Generation of chemically competent *E. coli*

To generate chemically competent *E. coli,* 50 ml of SOB medium were inoculated with a small amount of bacteria from a previous batch of competent cells. This culture was grown overnight and subsequently used to inoculate 200 ml of SOB. Once grown to an OD_{600} of 0.5, cells were harvested by centrifugation at 4,500 rpm for 10 min at 4°C. The bacterial pellet was subsequently carefully resuspended in 15 ml Tfb I buffer and rested on ice for 20 min. Cells were again pelleted and the pellet was resuspended in 4 ml of Tfb II buffer. After 10 min incubation on ice, the competent cell were distributed in 25 µl aliquots, flash frozen in N_2(l), and stored at -80°C

3.2.10 Generation of ultra-competent *E. coli*

In order to generate ultracompetent expression cells, suitable for double transformations and co-expression of multiple proteins from two vectors, 25 ml of SOB medium were incubated with one aliquot of chemically competent *E. coli* BL21 (DE3) RIL. Cells were grown for 6 – 8 h at 37°C shaking at 250 – 300 rpm. 125 ml of SOB medium were then inoculated with 0.5 ml of the pre-culture and incubated overnight at 18°C. If OD_{600} exceeded 0.55 in the morning, 10 ml of the overnight culture were transferred to 125 ml of fresh SOB medium, which was subsequently grown to OD_{600} = 0.55.

Reaching a density of 0.55, cells were harvested at 2,500 x g for 10 min at 4°C. The supernatant was removed completely before the cell pellet was gently resuspended in 40 ml of ice-cold Inoue - transformation buffer by gentle swirling. The centrifugation step was repeated as described above, this time resuspending the pellet in 10 ml of transformation buffer afterwards. After addition of 50 µl of DMSO the cell suspension was carefully mixed and incubated on ice for 10 min. Finally, the ultracompetent *E. coli* were aliquoted and flash frozen in N_2(l).

3.2.11 Transformation of chemically (ultra) competent *E. coli*

XL1blue or BL21 RIL strains were used for cloning or expression of proteins respectively. The chemically competent cells were first thawed on ice for 10 min, then ~50 ng plasmid were added before letting cells rest on ice for another 25 min. After a heat shock at 42°C for 40 s, cells were put on ice for another 5 min, then 500 µl of room temperature LB medium were added. Afterwards, cells were generally cured for an hour at 37°C under vigorous shaking. Agarplates containing suitable antibiotics were used to select for positively transformed clones.

3.3 Expression, purification, and optimization of recombinant proteins

3.3.1 In silico structure prediction

During this project, a number of freely available online tools were used to predict secondary structure elements and to identify potential domains within each of the UAF subunits. The prediction of secondary structure elements, such as helices, beta sheets, and disordered regions, was carried out using the combined outputs of Quick2D and PsiPred (including DisoPred). HHpred was employed to predict domains and determine their boundaries for each subunit. The respective primary amino acid sequences were analyzed independently, and the output of each tool was compared to predict these elements with the highest possible accuracy. This information was utilized to design and optimize constructs beyond the scope of full - length proteins. Furthermore, AlphaFold was employed to predict 3D structures in the later stages of this work.

3.3.2 Iterative protein analysis and construct optimization

To optimize stability and enhance the purity of each construct for crystallization and functional assays, purification products were comprehensively analyzed. SDS - PAGE of the final products served as an indicator of purity and was complemented by α His western blots to confirm potential degradation bands as the respective target proteins. Analytical size exclusion chromatography runs were repeated after 24 hours at 4°C to observe time - dependent degradation or multimerization.

3.3.3 Expression and purification of Uaf30

Full - length Uaf30 was expressed in *E. coli* BL21 (DE3) RIL from a pCDFduet-1 vector with an N terminal 6xHis-tag. Expression was carried out in LB medium at 37 °C for 3 h after induction at OD_{600} = 0.5 with 0.2 mM IPTG. Lysis was achieved using a Branson Sonifier (output control: 5 / 80% duty cycle) for 5 x 5 min with Uaf30 lysis buffer. The whole cell lysate was cleared via centrifugation at 20,000 rpm (~48,000 *g) for 45 min at 4 °C. Uaf30 was captured using a 5 ml HisTrap column on an ÄKTA™ pure 25 chromatography system. The

column was washed with 5 CVs of Uaf30 lysis buffer to remove excess lysate and unbound protein. In order to decrease the nucleic acid content of the bound sample and reduce the amount of unspecifically bound proteins on the column, a high salt (Buffer Uaf30 HS) and an imidazole washing step (Uaf30 W75) were performed. Elution was done in Uaf30 elution buffer at 350 mM imidazole. After examining wash and elution fractions via SDS-PAGE, fractions containing Uaf30 were pooled, concentrated, and further purified using a Superdex 75 Increase 10/300 GL size exclusion chromatography column using Uaf30 SEC buffer. The purest Uaf30 - containing fractions were isolated. Aliquots were flash frozen and stored at -80°C.

3.3.4 Expression and purification of Uaf30 N & C-terminal domain

The two domains Uaf30 Ntd (1-61) and Ctd (115-195) were both expressed in *E. coli* BL21 (DE3) RIL from a pOPIN - B vector with an N terminal 6 x His tag. Expression was carried out at 18 °C overnight in LB medium after inducing expression at OD_{600} = 0.5 with 0.2 mM IPTG. All steps of the Ni^{2+} affinity purification were done as described in the previous section (Purification of full length Uaf30). Uaf30 Ntd was further purified using a HiTrap Heparin column. For this, all fractions containing Uaf30 Ntd were pooled and loaded onto the column at 200 mM NaCl (Uaf30 N Hep 200 buffer). After a washing step at 300 mM NaCl, the protein was eluted at 500 mM NaCl (Uaf30 N Hep 500 buffer). Both proteins were finally isolated via size exclusion chromatography using a Superdex 75 Increase 10/300 GL column and Uaf30 SEC buffer.

3.3.5 Expression and purification of the Rrn9 / Rrn10 subcomplex

Rrn9 and Rrn10 were co-expressed in *E.coli* BL21 (DE3) RIL cells from a pOPIN-B plasmid containing an N-terminal 6x His Tag fused to Rrn9. Expression was carried out in LB medium at 18°C overnight after induction at OD_{600} = 0.5 with 0.2 mM IPTG. Lysis was achieved using a Branson Sonifier (Output control: 5 / 80% duty cycle) for 5 x 5 min with Rrn9/10 lysis buffer. The whole cell lysate was cleared via centrifugation at 20,000 rpm (~48,000 *g) for 45 min. His - Rrn9/10 was captured with a 5 ml HisTrap column on an ÄKTA™ pure 25 chromatography system. The column was first washed with 5 CV of Rrn9/10 lysis buffer to

remove excess lysate and unbound proteins. To reduce the amount of unspecifically bound proteins, the column was washed with 5 CV of Rrn9/10 W75 buffer. Elution was done in Rrn9/10 elution buffer at 350 mM imidazole. After analyzing wash and elution fractions via SDS-PAGE, fractions containing Rrn9/10 were pooled, concentrated and further purified using a Superdex 200 Increase 10/300 GL size exclusion chromatography column with Rrn9/10 SEC buffer. Due to strong oligomerization over time and high yield of this subcomplex, monomeric fractions could not be properly isolated in the first SEC run without strongly diluting the sample. For this reason, fractions with the least possible amount of degradation were pooled and immediately applied to a size exclusion a second time under the same conditions, eliminating the void peak and separating the monomer peak more clearly from the multimers. Again, fractions showing the least degradation of Rrn9 in SDS PAGE were promptly aliquoted after elution and flash frozen for storage at -80 °C.

3.3.6 Tag cleavage for crystallization batches

Tag cleavage was done for protein batches that were subsequently used for crystallization screenings. This step was conducted after the first affinity purification step. Due to inactivity issues of HRV 3C in presence of imidazole, a buffer exchange was done for the elution fractions of the His-Trap columns to prepare the protein for tag cleavage. For this purpose, a PD 10 column (GE Healthcare) was performed according to the manufacturers protocol, using IEX200 - binding buffer for application and elution. The tag was cleaved off using lab own stock of HRV 3C protease in a ratio of 1:1000 (w/w). Protease and tag were then removed from the protein via reverse Ni^{2+} affinity purification. Prior to the application to the Ni^{2+} affinity column, Imidazole was spiked into the eluted fractions to reach a final concentration of 20 mM to avoid unspecific binding of the proteins of interest to the column.

3.3.7 Expression and purification of UAF and UAF mutants

Except for minor changes, the expression and purification for the whole UAF complex was inspired by an existing protocol by Smith et. al (2018). All six subunits the UAF complexes were co-expressed from a pCDFduet-1 and a pOPIN-B vector, each coding for three subunits. Co-transformation of both vectors was achieved using ultra competent *E. coli*. Except for the histone proteins H3 and H4, every other subunit was N terminally His tagged. Overexpression was done in *E. coli* BL21 (DE3) RIL using TB autoinduction medium supplemented with 150 ml of 20 x NPS and 50 ml of 5052 sugar mix for each 2 L culture. Cells were grown at 37 °C until a density of OD_{600} = 0.5 was reached. The cultures were cooled down on ice for 10 min. Subsequently, they were transferred to a 24°C shaking incubator and grown for another 18 h before the cells were pelleted at 4,000 rpm, flash frozen in N_2(l) and stored at -80°C.

Each 1 l pellet was resuspended in 60 ml UAF lysis buffer. Cells were lysed using a Branson sonifier (output control: 5 / 80% duty cycle) for 6 x 5 min. The whole cell lysate was cleared via centrifugation at 20,000 rpm (~48,000 *g) for 1 h. A 5 ml His-Trap column was utilized to capture the multi-his-tagged protein using an ÄKTA™ pure 25 chromatography system. After the sample was applied to the column, it was washed with 5 CV of UAF lysis buffer. The column was removed from the ÄKTA system for a subsequent ATP wash at RT. For this, 2.5 CV of ATP washing buffer (Buffer L supplemented with 5 mM ATP and denatured protein) were applied to the column. After 15 min incubation at RT the column was further washed with 2.5 CV of UAF ATP wash buffer before the ÄKTA program was continued. The column was subsequently washed with 5 CV of UAF wash buffer. Afterwards, the imidazole concentration was increased to 50 mM to reduce the amount of unspecifically bound protein. Elution was done at in UAF elution buffer at 350 mM imidazole. After analyzing wash and elution fractions via SDS-PAGE, UAF-containing fractions were pooled and loaded onto a 1 ml HiTrap SP HP column in UAF S load buffer at 350 mM KCl (diluted from 450 mM with salt free UAF elution buffer). 5 CV of UAF S 500 buffer were used to wash the bound sample before a step elution was done in UAF S 750 buffer to yield a pure UAF peak. All peak fractions were pooled and concentrated before the samples were further purified using a Superose 6 Increase 10/300 GL size exclusion column with UAF SEC at 450 mM KCl. Fractions containing whole UAF were identified via SDS PAGE and flash frozen for storage at -80°C.

3.3.8 Expression and purification of *C.t.* Uaf30-like

Full - length *c.t.* Uaf30-like was expressed in *E. coli* BL21 (DE3) RIL from a pOPINB vector with a cleavable N terminal 6 x His-tag. Expression was carried out in LB medium at 37 °C for 3 h after induction at OD_{600} = 0.5 with 0.2 mM IPTG. To reduce nucleic acid content of the final sample, DNase was added to the lysis buffer (25 µg/ml) and the solution was incubated on ice for 15 min prior to lysis. Lysis was achieved using a Branson Sonifier (output control: 5 / 80% duty cycle) for 5 x 5 min with Uaf30 lysis buffer. The whole cell lysate was cleared via centrifugation at 20,000 rpm (~48,000 *g) for 45 min at 4 °C. The protein was captured using a 5 ml HisTrap column on an ÄKTA™ pure 25 chromatography system. The column was washed with 5 CVs of Uaf30 lysis buffer to remove excess lysate and unbound protein. In order to further decrease the nucleic acid content of the bound sample and reduce the amount of unspecifically bound proteins on the column, a high salt (Buffer Uaf30 HS) and an imidazole washing step (Uaf30 W75) were performed, as described for the yeast homolog. Elution was done in Uaf30 elution buffer at 350 mM imidazole. After examining wash and elution fractions via SDS-PAGE, fractions containing *c.t.* Uaf30-like were pooled and the sample was appropiately diluted to 3.5 ml for the following PD 10 – desalting column (see manufacturers manual). The subsequent tag cleavage was conducted as described in *3.3.6.* After this process, the sample was concentrated to 0.5 – 1 ml and applied to a Superdex 200 Increase 10/300 GL size exclusion chromatography column, using Uaf30 SEC buffer. The purest *c.t.* Uaf30-like - containing fractions were isolated. Aliquots were flash frozen and stored at -80°C.

3.3.9 Expression and purification of *C.t.* Rrn5-like Ntd

The N terminal domain of *c.t.* Rrn5-like was expressed in *E. coli* BL21 (DE3) RIL from a pOPINB vector with a cleavable N terminal 6 x His-tag. Expression was carried out in LB medium at 37 °C for 3 h after induction at OD_{600} = 0.5 with 0.2 mM IPTG. Lysis was achieved using a Branson Sonifier (output control: 5 / 80% duty cycle) for 5 x 5 min with Uaf30 lysis buffer. The whole cell lysate was cleared via centrifugation at 20,000 rpm (~48,000 *g) for 45 min at 4 °C. The protein was captured using a 5 ml HisTrap column on an ÄKTA™ pure 25 chromatography system. The column was washed with 5 CVs of Rrn5-like lysis buffer to remove excess lysate and unbound protein. In order to reduce the amount of unspecifically

bound proteins on the column, an imidazole washing step (Rrn5-like W75) were performed. Elution was done in Rrn5-like elution buffer at 350 mM imidazole. After examining wash and elution fractions via SDS-PAGE, fractions containing *c.t.* Rrn5-like were pooled and the sample was appropiately diluted to 3.5 ml for the following PD 10 – desalting column (see manufacturers manual). Tag cleavage was done as described in *3.3.6.* After this, the sample was concentrated to 0.5 – 1 ml and applied to a Superdex 75 Increase 10/300 GL size exclusion chromatography column, using Rrn5-like SEC buffer. The purest *c.t.* Rrn5-like - containing fractions were isolated. Aliquots were flash frozen and stored at -80°C.

3.4 Biochemical methods

3.4.1 SDS PAGE

Protein samples that had to be analyzed via SDS-PAGE were mixed with 5x SDS loading dye. Samples were boiled for 10 min at 95°C. Gels were run for 25 - 35 min at 200 V depending on the molecular weight range that had to be visualized. Commercially bought gels (12 % Bis-Tris gels or 4 – 12 % Bis-Tris gradient gels) were run at 170 V for 45 – 60 min. Subsequent staining was achieved using Instant Blue ™ coomassie staining solution.

3.4.2 Mass Spectrometry

MS measurements were outsourced to the lab of Dr. Astrid Bruckmann (University Regensburg). Sample preparation for Mass spectrometry followed all purification steps described in the purification chapters. Commercially purchased SDS gels were used to decrease contamination risk. Bands of interest were cut out and analyzed by colleagues of the Bruckmann lab.

3.4.3 Western Blot

In order to detect His-tagged protein, α-His western blots (WBs) were performed. For this purpose, SDS-PAGE was first done as described in 3.4.1. Proteins were then transferred onto a PVDF membrane, using a BioRad Trans-Blot® Turbo ™ blotting system. For this, after

activation of the membrane by washing in methanol for 30 s, all components were soaked in WB transfer buffer. The gel was then stacked on top of the membrane, and both were placed between two triple stacks of Whatman paper. Blotting of the proteins was performed at 1,3 A and 25 V for 14 – 20 min. The membrane was then blocked in WB blocking solution for 1 h at room temperature or overnight at 4°C. Following this, the membrane was washed for 3 x 5 min with TBS-T. His Tags were generally detected using the HisProbe™ - HRP conjugate antibody. Binding was done for 1 h at RT in WB antibody binding buffer with a final antibody dilution of 1 : 5000. Afterwards, the membrane was washed again before using BM chemiluminescence blotting substrate (POD) and a LAS-3000 imager to detect the chemiluminescence. Signals were detected at high sensitivity, stacking 10 s increments.

3.4.4 Mass photometry

Mass photometry was done to screen the size distribution and oligomerization state of some purification products Prior to testing a sample, the device was calibrated to the proper molecular weight range using BSA, IgG and Thyroglobulin. The dilution or SEC buffer was first measured to focus the lens. Proteins were then added to the buffer drop in different ratios, diluting the protein to an appropriate concentration yielding a count of 3000 – 7000 (~ 5 - 10 nM) for measurements. Protein sizes were then measured for 1 min. The size distribution was plotted against the molecule count. To minimize the impact of protein dilution on the state of oligomerization, high protein concentrations, ranging up to ~50 µM, were maintained until shortly before application onto the lens of the photometer, matching the highest used concentrations of these proteins in functional and structural assays.

3.4.5 Protein quantification

Protein concentrations were usually measured using a NanoDrop One (Thermo Fisher) with the respective predicted Abs (1%) values predicted by 'Expasy ProtParam' at 280 nm. In rare cases, a Bradford assay was performed (Bradford 1976). A calibration line was prepared using increasing concentrations of BSA as a reference absorbant. Proteins of unknown concentration were diluted appropriately. All samples were mixed with 1x Bradford solution to a final volume of 1 ml. Extinction was measured at 595 nm wavelength.

3.5 X-ray crystallography

3.5.1 Pre crystallization tests

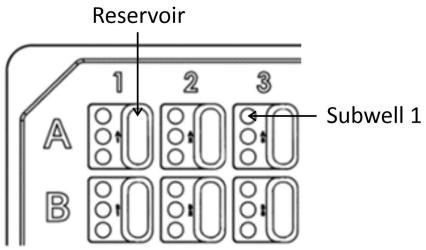
In order to determine a suitable protein concentration range for the subsequent screening process, each protein was first examined in a pre-crystallization test (PCT). For this, drop ratios (buffer : protein) of 0.5 µl : 1 µl, 1 µl : 1 µl, and 1 µl : 0.5 µl were set with the provided A1, A2, B1, and B2 solutions from the PCT kit (Hampton Research), starting with the highest possible protein concentration after size exclusion and the centrifugal concentration process. Drops were quickly set onto glass cover slides before they were placed top down (hanging drop vapor diffusion) onto pre - greased 24 well deep well plates with above mentioned solutions as reservoir (500 µl). The plates were stored in a fridge at constant 20°C and inspected regularly over the next two days. According to the kits protocol, protein concentrations for the initial high throughput (HT) screenings were adapted depending on the morphology of the precipitate in the different conditions of the PCT.

3.5.2 High throughput (HT) crystallization screening

For the initial HT screening process, a variety of buffer screens were used for all of the constructs. Unlike for the pre-crystallization test, sitting drop vapor diffusion was chosen for these screenings. A mosquito® LCP robot (SPT Labtech) was used to set drops of 300 nl (150 nl protein + 150 nl reservoir solution) into one to two of three wells of Intelli-Plate 96 – 3 LVR (low volume reservoir) plates *(Figure 6)*.

Figure 6: X-tal Intelliplate setup scheme.

Schematic representation of Intelliplate subwells used for initial crystallization screenings using sitting drop vapor diffusion. The reservoir was filled with mother liquor; subwell 1 and 3 were used for protein, subwell 2 for buffer controls. Each drop contained 300 nl of mother liquor and 300 nl of protein or size exclusion buffer respectively.



The second subwell was generally used as a buffer control well, serving as a reference drop with the respective SEC buffer instead of protein. The plates were tightly sealed with transparent qPCR plate sealing foils and stored in a tempered room (20°C) in the dark. A variety of protein concentrations were tested for the purified constructs. Plates were inspected under a high magnification light microscope in regular time intervals (day1, day2, day 3, day 7, ...) spanning the following three months after setting the drops.

3.5.3 Finescreening

Hits in the initial high throughput screening process were followed up with finescreens. Finescreens were generally done in 24 well formats with hanging drop vapor diffusion. Drops were set up as described in the PCT chapter. Conditions of the finescreens were chosen based on the initial hit condition of the buffer screen. Two variables, e.g., precipitate concentration, salt concentration, or pH were decreased and increased in small increments, resulting in a fine grid screen around the initial condition. The resulting buffers were used as reservoir solution and mixed in different ratios (see PCT) with the respective protein using the same concentration that was initially used in the screening process. Plates were stored in a fridge at constant 20°C and inspected regularly over the next months.

3.5.4 Crystal handling / Freezing

Drops containing crystals were slowly equilibrated by adding small volumes of the respective freezing buffer, containing additional PEG or glycerol, twice over a time period of 15 min. Crystals were then carefully fished using loops of appropriate size and transferred into N_2(l) and stored in a liquid nitrogen tank.

3.6 Functional Analyses

3.6.1 Electromobility Shift Assays (EMSAs)

3.6.1.1 Standard EMSAs

Standard EMSAs were performed using the wildtype promoter template ranging from -183 / +8 relative to the transcription start site. The control template contained a randomized upstream element (UE) (randomized region: -155 / -38). 5 nM (50 fmol) of Cy5 labelled DNA template were incubated with increasing amounts of protein. Reactions were conducted in 20 mM Tris/HCl pH 7.8, 10 % glycerol, 1 mM $MgCl_2$, 0.1 mg/ml BSA, and 1 mM DTT (2 µl 5x EMSA Reaction buffer per reaction) in a final volume of 10 µl per reaction at 200 mM salt concentration. After 30 min incubation at RT, 2.25 µl OrangeG loading dye were added and loaded onto pre-run 6 % polyacrylamide gels (30 min / 115 V). Gels were then run for 60 – 80 min at 115 V in 0.4 x TBE. The Cy5 signal was subsequently detected using a Typhoon FLA 9500 imager using the Cy5 channel.

3.6.1.2 Promoterfragment EMSAs

Wild type rDNA promoter fragments (*S. cerevisiae*) of 40 bp length ranging from -140/-100 (#1), -110/-70 (#2), -80/-40 (#3) and -50/-10 (#4) (relative to the TSS) were annealed using complementary single-stranded oligos. For this, they were equimolarly mixed in a volume of 50 µl in DNA annealing buffer, heated up to 98 °C for 10 min in a PCR cycler, then the temperature was lowered by 1 °C / min down to 15 °C.

50 nM (500 fmol) of each fragment were incubated with increasing amounts of Uaf30 or Rrn9/10 (50 nM – 5 µM). The reaction was conducted in 20 mM Tris-HCl pH 7.8, 10 % glycerol, 1 mM $MgCl_2$, 0.1 mg/ml BSA, and 1 mM DTT in a final volume of 10 µl per reaction at 200 mM salt concentration. KCl was supplemented according to the salt concentration of the respective SEC buffers of the proteins to match 200 mM. Reactions were incubated at RT for 30 min. 2.25 µl 6x loading dye (OrangeG) were added before 12 µl of each reaction were loaded onto a pre-run 10% native acrylamide gel and run for 85 min at 115 V in 0.4 x TBE.

Staining was achieved by bathing the gels in 0.4 x TBE + 1: 20,000 SYBR Safe for 5 min. Images were taken on a Typhoon FLA 9500 using the SYBR Safe Channel, detecting emission at 473 nm.

The amount of unbound DNA was quantified using FIJI (ImageJ). All bands showing unbound DNA were therefor equally captured and the signal intensities were plotted. Integrals of each band were measured, and the values of each data point (increasing protein concentrations) were normalized to each respective negative control (no protein – unbound DNA only). At least 5 replicates per data point were measured for each fragment to reduce the impact of pipetting errors etc.. Replicates for the pairs 375 nM + 625 nM and 500 nM + 750 nM were done in seperate runs, as space on each gel was limited by usage of 10 well combs to ensure sharp bands for quantification. Mean values from all runs were calculated and plotted against the protein concentration.

Statistical analysis:

To identify statistically significant outliers, multiple comparison tests (Tukey contrast; T-tests) were conducted for every pair (1 – 2 / 1 – 3 / 1 – 4 / 2 – 3 / 2 – 4 / 3 – 4) within every group (protein concentration). Statistical evaluation was done with the help of Maximilian Pichler.

3.6.1.3 Competetive EMSAs

Competing fragments of the *S. cerevisiae* rDNA promoter were PCR amplified from two different vectors containing the wild type rDNA promoter sequence and an rDNA promoter sequence with a randomized upstream element ('rUE' – randomized from -155 / -38). The amplified area ranged from -212 to +119 relative to the transcription start site. A Cy3-labelled oligo was used to amplify the WT-fragment while the rUE-fragment was labelled with Cy5. A final concentration of 5 nM of each template was incubated with increasing amounts of each respective protein. Sample preparation and incubation were done as described in *3.6.1.2.* To account for the increased template (and protein) sizes, 6 % polyacrylamide gels were used and were run for 2:15 h at 115 V.

Cy5 and Cy3 channels were scanned separately using a Typhoon™ FLA 9500 laser scanner. In order to equalize signal intensities of the two channels, the intensity of both negative

controls were matched in FIJI (ImageJ) before the images were merged with magenta representing the wt– and yellow the rUE– template. Bright pixel outliers were removed with a threshold of 50% contrast and a maximum pixel size of 8 px.

3.6.2 Fluorescence Anisotropy

Fluorescence anisotropy (FA) is a sensitive photometric method, used to determine changes in physical properties of a fluorescently labeled component, in this case: Cy-5 labelled DNA. With binindg of the protein, the size of the complex increases, directly affecting its movement patterns (for a more detailed description refer to (Gijsbers et al. 2016; Trabulo et al. 2012). DNA-fragment annealing, sample preparation, and reaction conditions were the same as described in section *3.6.1.2*. 10 nM of Cy5 labelled fragments #1 – #4 were incubated with increasing amounts of the respective protein (16 data points ranging from 5 – 100,000 nM). Reactions and measurements of all four fragments were generally done in parallel from the same protein dilution series to reduce pipetting and batch variations between the fragments in each replicate. After 30 min of incubation at 25 °C, all reactions were transferred into black 384 well round bottom microtiter plates (Corning®). Anisotropy measurements were done on a TECAN Spark® (Cy-5 Channel, G-Factor = 1). For each replicate, 10 cycles were measured and averaged for each reaction to yield the mean anisotropy values. Three replicates were done for each fragment and from different protein dilution series and purification batches. Anisotropy values of the three replicates were averaged and plotted against the protein concentrations. Curve fitting was done in MatLab with protein concentration in both log10 and linear scale. Assay and data fitting were done with the help of David Stelzig from Prof. Dr. Remco Sprangers group. The final fitting and formatting script can be found in section *6.1.7*.

3.6.3 In vitro transcription

To assess the transcriptional activity of UAF and evaluate the effects of deletions on its activity, we conducted in vitro transcription assays. Transcription rates were measured using two distinct templates, both of which start at position -212 relative to the TSS. The wildtype fragment extended up to position +119, while the control template (containing a randomized

upstream element from -155 to -38) extended up to +147. This resulted in RNA transcripts of 119 nucleotides (wt) and 147 nucleotides (rUE) in length, respectively.

0.125 pmol of both DNA templates were first incubated together with 0.5 µl of Net1 – C and 5x Trx Buffer without NTPs. The salt concentration was adjusted to match 200 mM after the following addition of UAF, using 1 M KCl and Trx-Buffer H0 in a total volume of 4 µl. After 5 min at RT, 0.5 pmol of UAF or the deletion mutants were added to the reaction. This assembly was incubated for another 20 min at RT before 0.25 pmol of Core Factor and 1 pmol of TBP in a total volume of 1.5 µl were supplemented to each reaction. After 1:20 h at RT, 12.5 µl of pre-incubated Pol I – Rrn3 (0.125 µM of Pol I pre-incubated with 1.75 µM of Rrn3 (1:14) prior adjusted to 200 mM salt concentration using KAc – dilution buffer and H0 – buffer) were added and assembled for another 30 min at RT. Transcription was started by addition of 5x Trx Buffer supplemented with 1 mM of ATP, GTP, UTP, 0.05 mM CTP, and 0.3 µCi of α^{32}P-CTP. The reaction was stopped after 30 min by adding 200 µl of Proteinase K Buffer (containing 0.5 mg/ml Proteinase K, 100 ng/µl Glycogen). Proteins were then digested at 56°C for 15 min. 700 µl of ice cold ethanol (100 %) were added and the RNA transcripts were precipitated over night at -20°C. The RNA was then pelleted by centrifugation at 13,000 rpm, 4°C for 15 min. The ethanol was carefully removed, and the RNA pellet was washed with 165 µl ice cold ethanol and centrifuged again for 10 min. Then, the supernatant was discarded, and the pellet was dried on a heating block for 30 seconds at 95°C. The dried RNA pellet was resuspended in 15 µl RNA loading dye. An 8 M UREA 6 % polyacrylamide gel was used to separate the transcripts. This gel was pre-run for 30 min at 25 W and, after sample application, run for another 35 min to separate the RNA. Whatman papers were utilized to transfer gels from the glass plates into water, the gel was carefully washed for 2 min and subsequently dried on a vacuum heater plate before the radioactive signal was detected overnight. The signals were visualized using a Molecular Imager ® PharosFX ™ Plus imaging system.

The protocol was established by Michael Pilsl, who also kindly provided all necessary transcription factors used in this assay, as well as Pol I, preincubated with Rrn3 (Pilsl 2021).

4 Results

4.1 Construct overview – *S. cerevisiae*

Over the course of this project, numerous proteins were examined. *Figure 7* gives an overview of all constructs that were either subjected to high throughput crystallization screenings or functional analysis. Deriving from full length UAF subunits, deletion mutants and fragments / domains were generated to improve stability, reduce flexibility, and increase purity for functional and structural characterization. These improvements were guided by combined insights from secondary structure and domain predictions extracted from Quick2D, PsiPred, or HHpred and AlphaFold, respectively. The computational predictions were supplemented with findings from in vitro protein analysis, e.g. SDS-PAGE, Western blot, and mass spectrometry.

HHpred domain predictions for all subunits, except the histone proteins H3 and H4, were done several times over the period of this thesis. First, domains could only be identified for subunits Uaf30 and Rrn5, while none could be mapped to Rrn9 or Rrn10. Two domains were identified for Uaf30: a winged helix motif, covering the first ~60 amino acids (1-61) and a SWI/SNF or MDM2 motif in its C terminal region, ranging from position ~115 to ~195. A SANT domain and a histone fold were found for Rrn5, mapping to its N (50 – 110) and C (195 – 275) terminal region respectively. *Figure 7* updated domain boundaries for the C terminal histone fold domain of Rrn5, that are based on the recently published model of UAF (160 – 250)(Baudin et al. 2022). In this model, the SANT domain was not reported. The histone folds of Rrn9 and Rrn10 that are mapped in *Figure 7* could not be identified until early 2022, when the domains were annotated in context of the publication of the UAF structure by Baudin et al., and are hence not included in the process of generating new constructs or their optimization in this work.

The domain boundaries were used to create several constructs for both subunits, of which only the expression of Uaf30 - Ntd and Uaf30 - Ctd yielded protein amounts that were sufficient for subsequent analysis. All C and N terminal truncations of Rrn9 and Rrn10 originate from in vitro analysis of purified Rrn9/Rrn10 subcomplex that was done before the structure was published.

Constructs overview *S.cerevisiae*

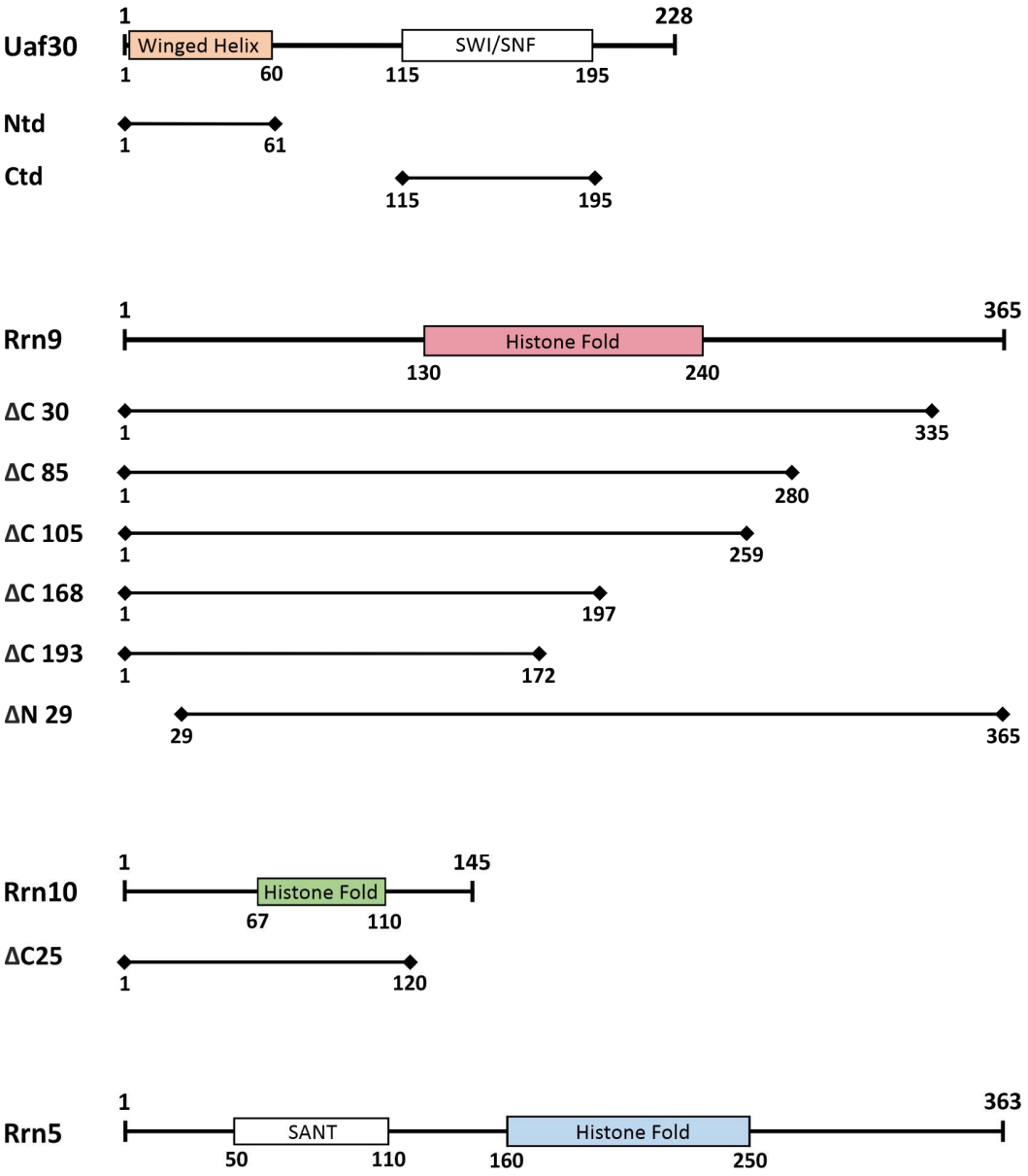


Figure 7: Construct overview *S. cerevisiae*.

An overview of all yeast UAF derived constructs that were used in this study. Colored domains were identified using HHpred or aligned and updated with the more precise boundaries from the recently published model of UAF (Baudin et al. 2022). Blank domains are not included in the publication, were however identified by HHpred. N and C terminal truncations, as well as domain boundaries (in numbers) include the position of the introduced stop codon. All truncations are named in respect to the amount of deleted amino acid on the respective terminus.

The occurrence of strong C terminal degradations was the driver for the generation of these C terminal truncations of Rrn9 as well as for Rrn10. Furthermore, structure predictions from Quick2D, PsiPred, and AlphaFold indicated a disordered N terminus of Rrn9, which inspired the creation of its ΔN truncation. Through site - directed mutagenesis, stop codons were introduced behind predicted secondary structure elements of Rrn9 and Rrn10, in front of the unstructured region in the C terminus of Rrn10, and behind the predicted disordered N terminus of Rrn9. The respective positions are highlighted in *Supplementary figure 4 & 5*.

Furthermore, Rrn5 and Uaf30 homologs from *Chaetomium thermophilum* yielded constructs that were exclusively tested in crystallization trials (refer to supplementary data: *6.1.1)*. Homologs were identified using NCBI blast, specifically targeting the *C. thermophilum* genome, as thermostability of proteins greatly improves chances of protein crystallization (Deller et al. 2016; Doerr 2006; Scandurra et al. 1998).

4.2 Expression and Purifications

4.2.1 Rrn9 and Rrn10 purify as a stable subcomplex

Co-expression and purification of Rrn9 in combination with Rrn10 was shown to increase the yield of either subunit in previous experiments done by myself and Christoph Engel (data not shown). As both subunits form a stable subcomplex, Rrn9 and Rrn10 were generally co-expressed from a single plasmid (pOPIN-B) fusing an N terminal 6 x His tag to Rrn9 and containing a short spacer with an ribosome binding site (RBS) sequence, separating both genes. Optimal expression of the subcomplex was achieved with expression over night at 18°C, using LB medium. Purification was done as described in *3.3.5.* In short, a 5 ml HisTrap was used to capture the stable Rrn9/10 complex. For all crystallization batches, the His tag was subsequently cleaved off using HRV - 3C protease after a buffer exchange via a PD10 column to remove imidazole, which was necessary for protease activity. A reverse HisTrap protocol was used to remove tag and protease. The final step for all batches were two consecutive size exclusion chromatography (SEC) runs in order to properly isolate the monomeric Rrn9/10 subcomplex from dimers and other aggregates.

Strong multimerization of the Rrn9/10 subcomplex was observed when running multiple SEC runs as the last step of purification. This effect could be observed already after a short time period (< 2 h) at 4°C *(Supplementary Figure 6)*. The previously mentioned re-runs of the isolated monomeric SEC fractions from the first runs were hence necessary to improve peak separation and minimize time dependent formation of multimers by flash freezing eluted fractions in N_2(l) immediately after peak elution during the second run. This process was also applied for any ΔN and ΔC truncations of the subcomplex. We were able to co – purify Rrn9 and Rrn10 as a stable subcomplex with a high yield of around ~4 – 7 mg per liter of culture.

The Rrn9/10 subcomplex additionally suffered C terminal degradation for both subunits *(Figure 8)*. This degradation could be observed throughout the whole purification process and did not improve when altering expression or purification protocols. Similar degradation patterns were observed in other studies working with Rrn9 aswell (Smith et al. 2018) (Siddiqi et al. 2001). The α His western blot (WB) indicated a C terminal degradation of Rrn9, as signals by the N terminal His tag were could be detected for all of the Rrn9 degradation bands. These bands were further analyzed via mass spectrometry (MS), yielding a peptide coverage that matches the indications of the WB *(Supplementary Figure 9)*. MS of the Rrn10 degradation also shows a clear decrease in peptide coverage for its C terminal region, which could not be detected via WB due to a lack of specific antibodies *(Supplementary Figure 10)*.

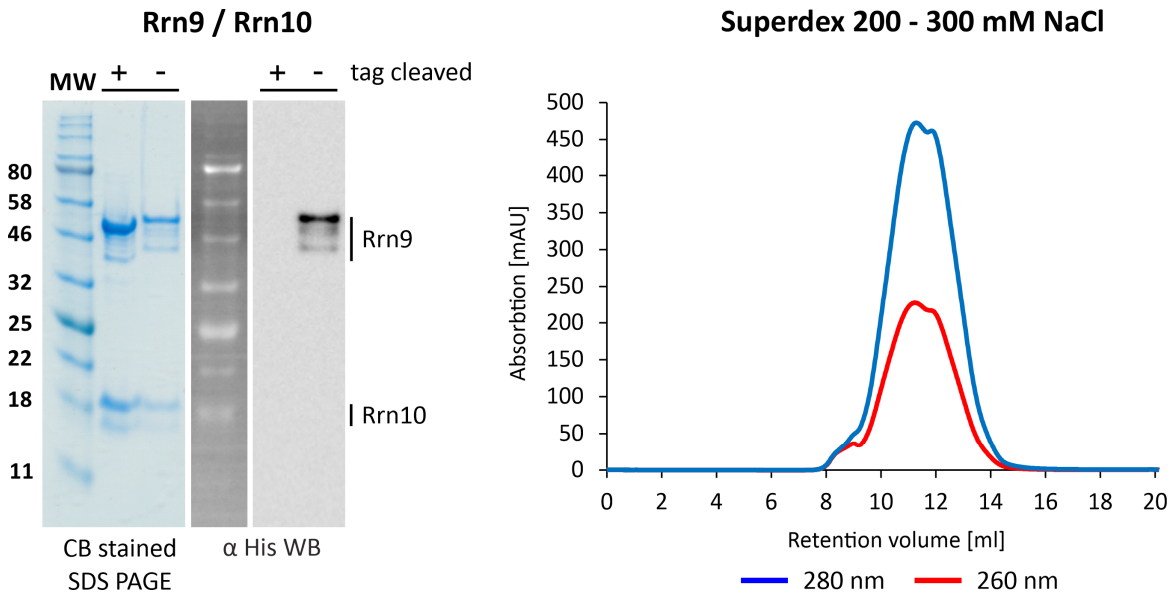


Figure 8: Purification of the Rrn9 / Rrn10 subcomplex (size exclusion).

(left) Coomassie stained SDS PAGE and α His Western Blot (WB) of the Rrn9/10 subcomplex with cleaved or uncleaved N terminal His tag (Lane 1 and 2 respectively). Degradation bands are visible below the main bands. For Rrn9, the western blot indicates C terminal degradation, as signals can be observed for the degradation pattern. Rrn10 is not tagged. **(right)** Size exclusion chromatogram (Superdex 200) of the Rrn9/10 subcomplex. The subcomplex elutes at a retention volume of 11.8 ml with visible heterogeneity (twin peak).

C terminal truncations were created in an attempt to stabilize the subcomplex and increase homogeneity of the sample for structural and functional analyses. These truncations led to a complex that showed less degradation with decreasing length of the C terminus of Rrn9 *(Figure 9 - gel 1),* with the most stable Rrn9/10 complex being a combination of a ΔC168 Rrn9 and a ΔC25 Rrn10 truncation *(gel 4).* Interestingly, the rate of dimer formation also seemed to decrease when Rrn9 loses its C terminus *(Supplementary Figure 6).* As expected, when comparing the truncations with the original degradation bands of Rrn9, ΔC30 and ΔC85 *(Lane 3 and 4 – gel 1)* were in a similar molecular weight range, while the longer truncations were significantly smaller. The ΔC25 Rrn10 truncation matched the size of its degradation band, indicating that this region is most likely degraded in the full length sample.

In addition to the C terminal truncations of Rrn9, an N terminal truncation was created. This ΔN29 variant was based on an Alphafold prediction of Rrn9 and the combined outputs of Quick2D and PsiPred, which also suggested a disordered N terminus. Co-purification of ΔN29 Rrn9 and ΔN25 Rrn10 yielded high amounts of protein that again showed C terminal degradations *(Figure 9 – gel 3)*.

From the purification of these different truncations we conclude that Rrn9 and Rrn10 form a stable subcomplex that is susceptible to degradation and dimerization. Additionally, from the ΔN29 and the ΔC193 truncation of Rrn9 we can narrow down an interface for stable interaction to position 30 – 172 of Rrn9 and position 1 – 120 of Rrn10. Furthermore, we conclude, that C terminal truncations enhance overall stability and reduce flexibility of the subcomplex.

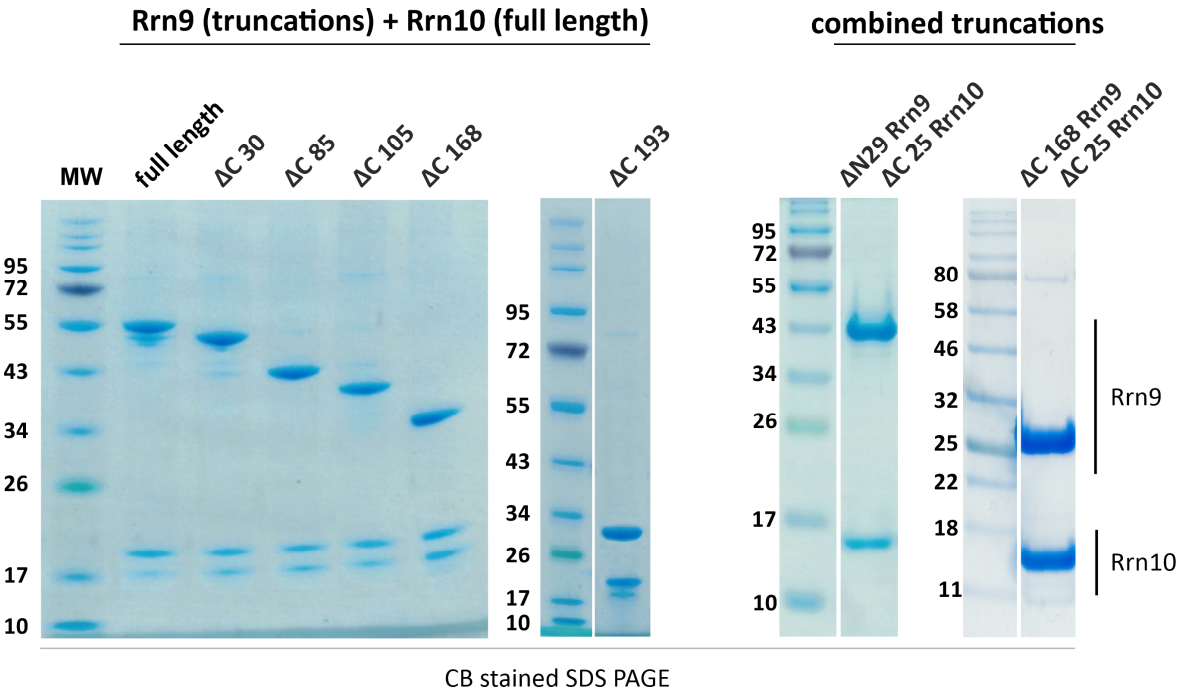


Figure 9: SDS PAGE of Rrn9 / Rrn10 truncations.

Coomassie blue stained SDS PAGEs of N and C terminal subcomplex truncations after size exclusion chromatography. Gel 1 and 2 show only C terminal Rrn9 truncations in combination with full length Rrn10. Gel 3 and 4 show combinations of N or C terminal Rrn9 truncations with a C terminal truncation of Rrn10. All variants co-purify as a stable complex, indicating that the minimal interaction interface is not disrupted. Note: different prestained molecular weight markers were used.

4.2.2 Purification of Uaf30 and its N & C terminal domain

Full length Uaf30 was expressed from a pCDFduet-1 vector, providing a non-cleavable 6 x His tag. Optimal expression was achieved within 3 h at 37°C after induction using LB medium. Purification was carried out as explained in *3.3.3.* In brief, a 5 ml HisTrap was used to capture Uaf30, followed by a SEC (Superdex 75) to isolate it from impurities.

Western blot analysis of the SEC product showed a faint band at roughly 22 kDa, indicating a neglectable C terminal degradation *(Figure 10)*. The additional band was analyzed via MS and the peptide coverage matched the indications of the western blot *(Supplementary Figure 11)*. With the aim to further improve the quality of the protein, two C terminal truncations were created (ΔC13 / ΔC33) *(Supplementary Figure 5)*. Interestingly, both showed highly reduced expression levels that were insufficient for structural or functional characterization.

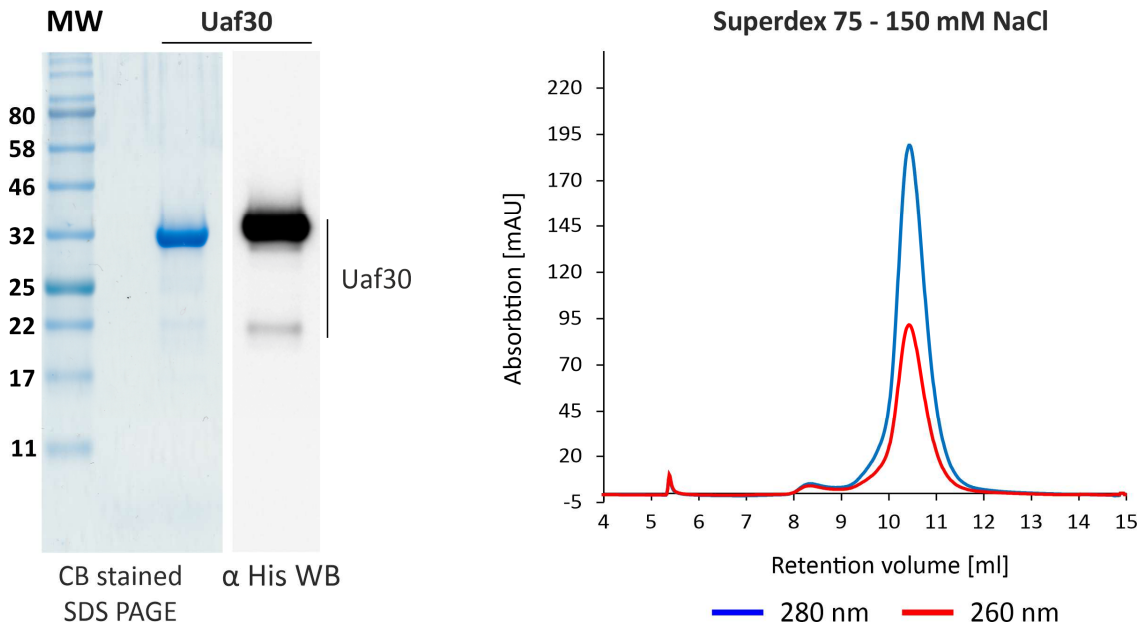


Figure 10: Purification of Uaf30 (size exclusion).

(left) Coomassie stained SDS PAGE and αHis Western Blot (WB) of Uaf30 with an N terminal His tag (non-cleavable). A faint degradation band is visible below the main bands. The western blot indicates C terminal degradation, as a signal is visible for the degradation band. **(right)** Size exclusion chromatogram (Superdex 75) of Uaf30. The subunit elutes at a retention volume of 10.5 ml.

As explained in the previous chapter, two domains could be identified for Uaf30 *(Figure 7)*. Expression of both domains was done from pOPIN-B vectors with N terminal His tags over night at 18°C in LB medium. Purification was achieved as described in *3.3.4*. Both domains were first subjected to a HisTrap. Uaf30-Ntd was then additionally purified using a Heparin column. Prior to a final SEC, tags were removed as previously described for Rrn9/10. Purification of both subunits yielded protein of very high purity *(Figure 11)*. While the Ctd could be concentrated to ~ 6 mg/ml, the N terminal domain had a solubility limit at ~ 1 mg/ml.

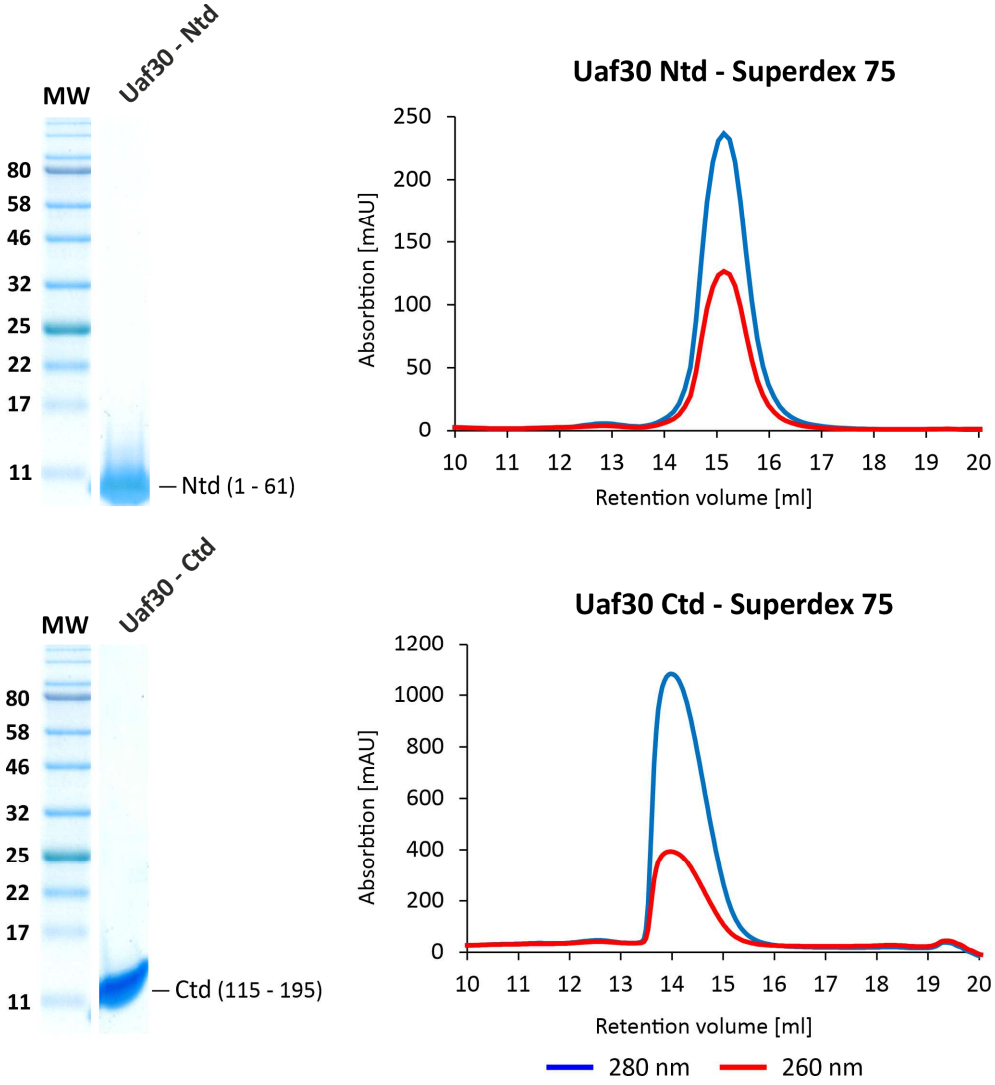


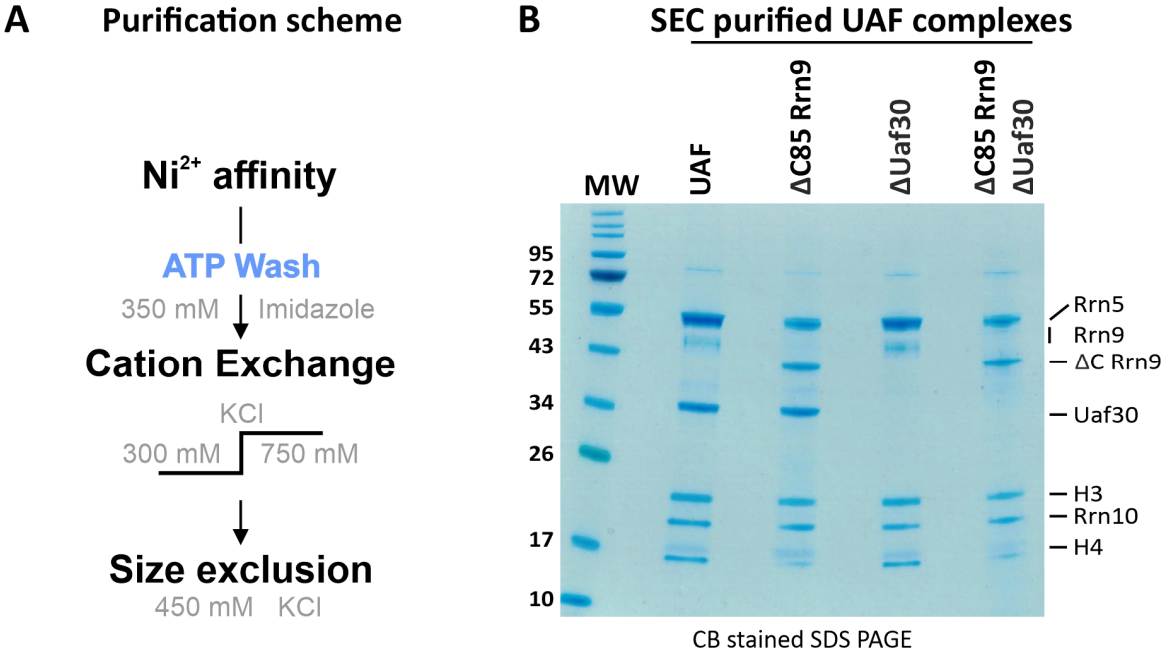
Figure 11: Purification of Uaf30 domains (size exclusion).

(left) Coomassie stained SDS PAGE of Uaf30 Ntd (upper left) and Ctd (lower left) with cleaved N terminal His tags. **(right)** Size exclusion chromatogram (Superdex 75) of Uaf30 Ntd and Ctd. The domains elute at retention volumes of 15 and 14 ml respectively.

4.2.3 Purification of UAF and three deletion variants

In addition to the structural and functional characterization of UAF subunits, domains, or truncations, the whole UAF complex and three deletion variants were examined. For the expression and purification of each complex, an *E. coli* expression and purification protocol from Smith et al. (2018) was used, with some minor adaptations *(see 3.3.7)*. In short, expression of all four complexes was done from two different vectors, carrying three subunits each. The initial vectors for expression of whole UAF were created with assistance of Antonia Neumeier (bachelor internship). Cells were grown over night at 25°C in autoinducing TB medium. After cell lysis, a HisTrap column with an additional ATP wash were performed to reduce Hsp70 contamination that was initially observed throughout purification of all UAF variants. The complexes were then further purified via cation exchange chromatography. A step wash and step elution were used to reduce volume and increase concentration of the fractioned sample before the following size exclusion (Superose 6).

The presence of all subunits was verified via SDS PAGE *(Figure 12 B)*. As mentioned above, in addition to the full length UAF complex *(Figure 12 B - Lane 2),* three mutants were generated to further investigate the functional relevance of the DNA binding domain of subunit Rrn9 and of subunit Uaf30. During the characterization of the Rrn9/Rrn10 subcomplex, the ΔC85 truncation of Rrn9, showed a loss of DNA binding in EMSAs (for reference see *4.3.2 / Figure 15*). Consequently, a UAF variant carrying this ΔC85 Rrn9 truncation was created *(Figure 12 B - Lane 3)*. Furthermore, based on the DNA binding properties of Uaf30 *(*for reference see *4.3.3 / Figure 16 - 18)*, a ΔUaf30 variant was created *(Figure 12 B - Lane 4)*. Both deletions were finally combined to a double deletion mutant (*Lane 5*). Consistent with our results from the purification of the Rrn9/10 subcomplex, faint degradation bands were detected in lane *2 and 4*, when Rrn9 was not C terminally truncated. Furthermore, subunit H4 seems slightly underrepresented in lane 3 and 5.


A
Purification scheme
Ni2+ affinity
ATP Wash
350 mM Imidazole
Cation Exchange
KCl
300 mM 750 mM
Size exclusion
450 mM KCl
B
SEC purified UAF complexes
MW
UAF
ΔC85 Rrn9
ΔUaf30
ΔC85 Rrn9 ΔUaf30
95
72
55
43
34
26
17
10
Rrn5
Rrn9
ΔC Rrn9
Uaf30
H3
Rrn10
H4
CB stained SDS PAGE

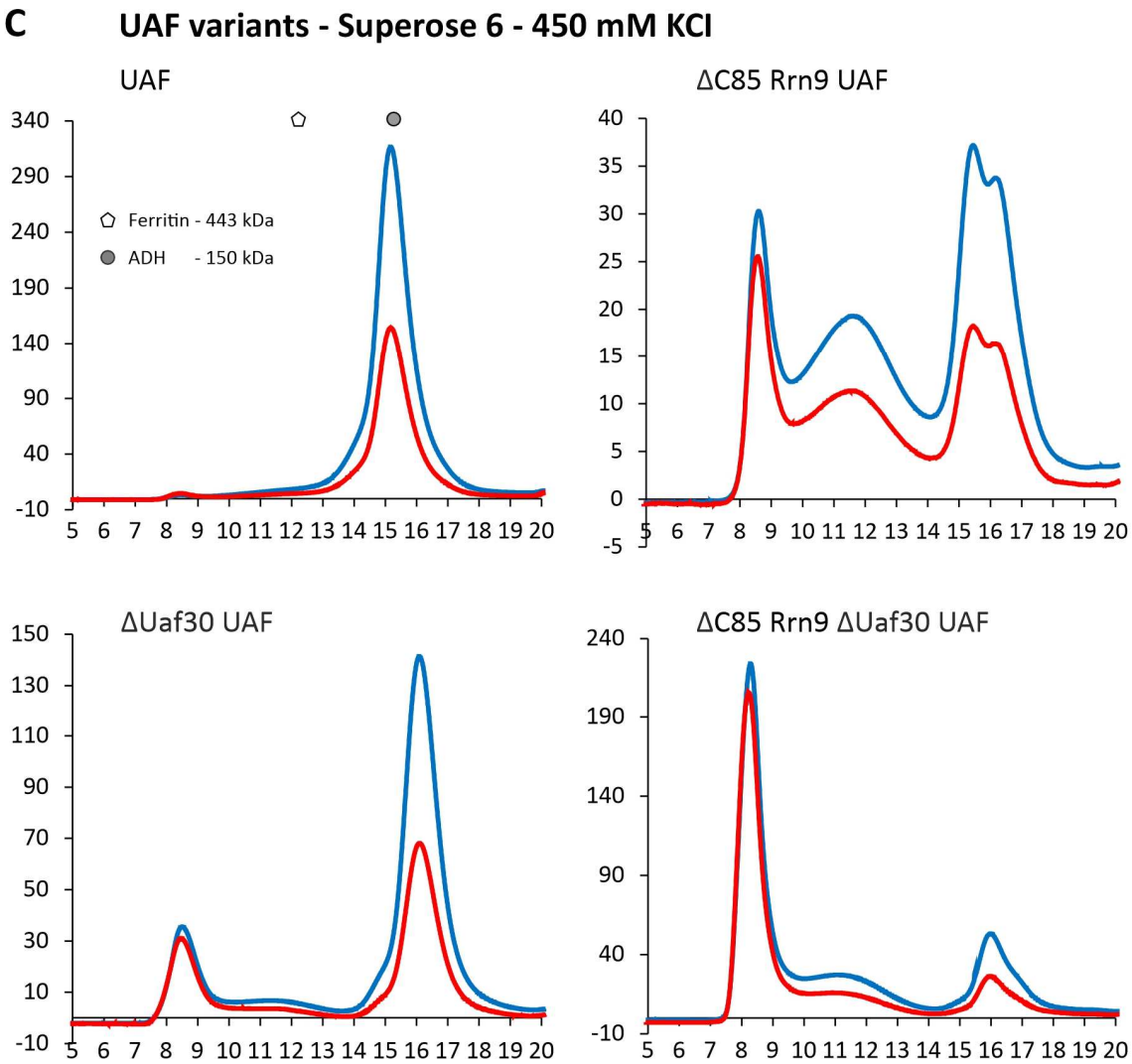

C
UAF variants - Superose 6 - 450 mM KCl
UAF
Ferritin - 443 kDa
ADH - 150 kDa
ΔC85 Rrn9 UAF
ΔUaf30 UAF
ΔC85 Rrn9 ΔUaf30 UAF

Figure 12: Purification of whole UAF and deletion mutants (size exclusion).

A) Purification scheme that was applied to all UAF variants. An ATP wash was conducted during the Ni^{2+} affinity chromatography to wash out chaperone contaminations (Hsp70). A cation exchange and a size exclusion were performed to further isolate all UAF variants. **B)** Comparative coomassie stained SDS PAGE of SEC fractions of whole UAF and three deletion mutants ΔC85 Rrn9, ΔUaf30 and the double - deletion mutant Δ C85 Rrn9 / ΔUaf30 (left to right). For comparison, equal amounts of all complexes were loaded onto the gel, confirming presence of all subunits and sucessful truncation/deletion in every complex. **C)** Size exclusion chromatograms (Superose 6) of the purified complexes. The purest fractions were pooled and concentrated. As a size reference, markers for the retention volumes of ferritin (443 kDa) and alcoholdehydrogenase (150 kDa) are included in the chromatogram of whole UAF (top left). Both reference proteins were run in the same buffer as all UAF complexes.

4.3 Functional characterization

4.3.1 Interaction of Rrn9 / Rrn10 with the rDNA promoter likely unspecific

To investigate whether the Rrn9/10 subcomplex specifically binds to the rDNA promoter, EMSAs were conducted using two distinct rDNA promoter templates: one containing the wildtype sequence and another containing a randomized upstream element (UE: -155 / -38) (for sequence alignment, refer to *Supplementary Figure 7)*. Both templates used in these standard EMSAs covered a length of 191 base pairs (-183 / +8). DNA binding was observed for each template within a similar protein concentration range (data not shown; refer to non-merged lanes in the competitive EMSA; *Figure 13*). For a more precise comparison of the affinities between both promoter variants, competitive EMSAs were performed using slightly longer rDNA promoter templates (-212 to +119). In these assays, the wildtype template was labeled with Cy3 (meganta), while the randomized UE template was labeled with Cy5 (yellow) *(Figure 13 A)*. No color transition was observed for the unbound DNA, as both templates were bound at the same nanomolar concentration range of 250 – 375 nM *(Figure 13 B)*. This suggests that Rrn9/10 not only binds DNA unspecifically but also lacks specificity for the wildtype sequence in this experimental setup.

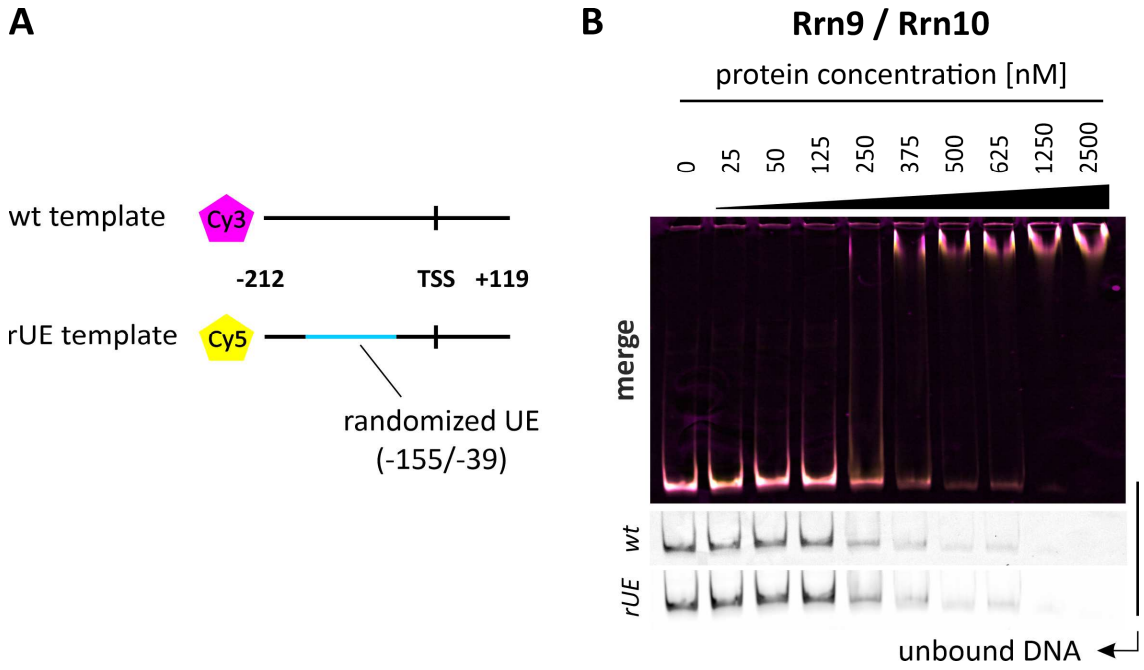


Figure 13: Competitive EMSA of the Rrn9 Rrn10 subcomplex.

A) Schematic depiction of the two competing DNA templates that were used for the competetive EMSAs. For direct sequence comparison see chapter 6.1.4. The Cy3 labelled template comprises the wildtype sequence stretching from position -212 to +119 relative to the transcription start site (TSS: position +1). The control sequence is randomized from position -155 to -39, substituting the whole upstream element (blue line). Sequence alignment is shown in Supplementary Figure 7 **B)** Competitive EMSA of the Rrn9 Rrn10 subcomplex, run on a a 6% native polyacrylamid-gel. The wildtype (Cy3) and randomized upstream element (Cy5) template are represented in magenta and yellow, respectively (as shown in A). Fluorescent signals were detected seperately (black/white channels) before they were merged in ImageJ. The absence of a color shift in the unbound DNA bands indicates that both fragments are bound with similar affinities.

To test this hypothesis, we performed promoter fragment EMSAs. This assay enabled us to narrow down potential sequence - specific interactions with one of four 40 bp wildtype promoter fragments, spanning the sequence from -140 to -10 relative to the transcription start site *(Figure 14 A)*. The amounts of unbound DNA from 5 replicates were averaged and the corresponding plots are shown *(Figure 14 B)*. Exemplary EMSAs are shown in *Supplementary Figure 8*. No distinct preference was evident at any given concentration, supporting our hypothesis that the Rrn9/10 subcomplex does not specifically bind a distinct sequence inside the promoter. Further testing involved the same 40 bp fragment approach in microscale thermophoresis (MST) and fluorescence anisotropy (FA). In both cases, the protein concentrations required for a reasonable K_d determination exceeded the solubility limit of the Rrn9/10 subcomplex (data not shown).

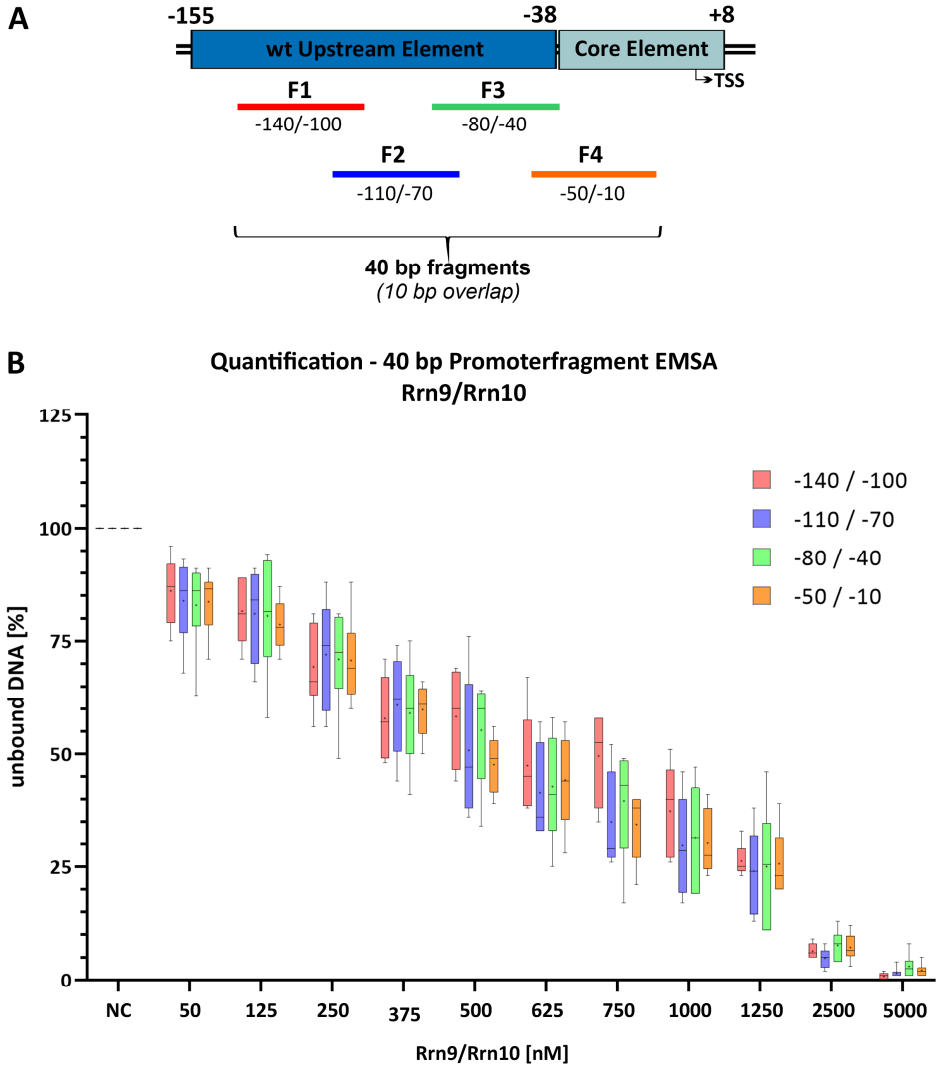


Figure 14: Quantification of Rrn9/10 promoterfragment EMSAs.

A) Schematic representation of wildtype promoter fragments that were used for promoterfragment EMSAs. 40 bp fragments with an overlap of 10 bp were annealed from single stranded oligos and span the entire upstream element and the core element, up to position -10, relative to the transcription start site. **B)** Unbound DNA from EMSAs of all four fragments quantified and plotted for every concentration. EMSAs and quantification of unbound DNA were conducted as described in *3.6.1.2.* Representative EMSAs are shown in *Supplementary Figure 8*. n ≥ 5 for all concentrations. Whiskers include all values (min to max), while the box extend from 25th to 75th percentile. The median value is indicated by a line inside the boxes, mean values by a +. No preference was detected for any of the DNA fragments, indicating a lack of specificity of the Rrn9/Rrn10 subcomplex.

4.3.2 DNA binding of Rrn9 relies on C terminal region

To determine what part of Rrn9/10 was responsible for DNA binding, we examined the DNA binding capability of the C terminal truncations of Rrn9 in complex with full - length Rrn10, following the same competitive EMSAs approach as for the full-length Rrn9/Rrn10 subcomplex. While the shortest truncation variant of Rrn9 (ΔC30) exhibited DNA binding *(Figure 15; left)*, introduction of a stop codon at position 280 (ΔC85) led to the loss of binding *(Figure 15; right)*. This outcome strongly suggests that the DNA binding by Rrn9 relies on its C terminal region. This conclusion is further supported by the observation that longer truncations of Rrn9 (for reference see *Figure 7)* also lost their ability to bind DNA (data not shown). Consequently, we can confidently map the location of a Rrn9 DNA binding domain to its C terminal region, spanning the amino acids from 280 to 335.

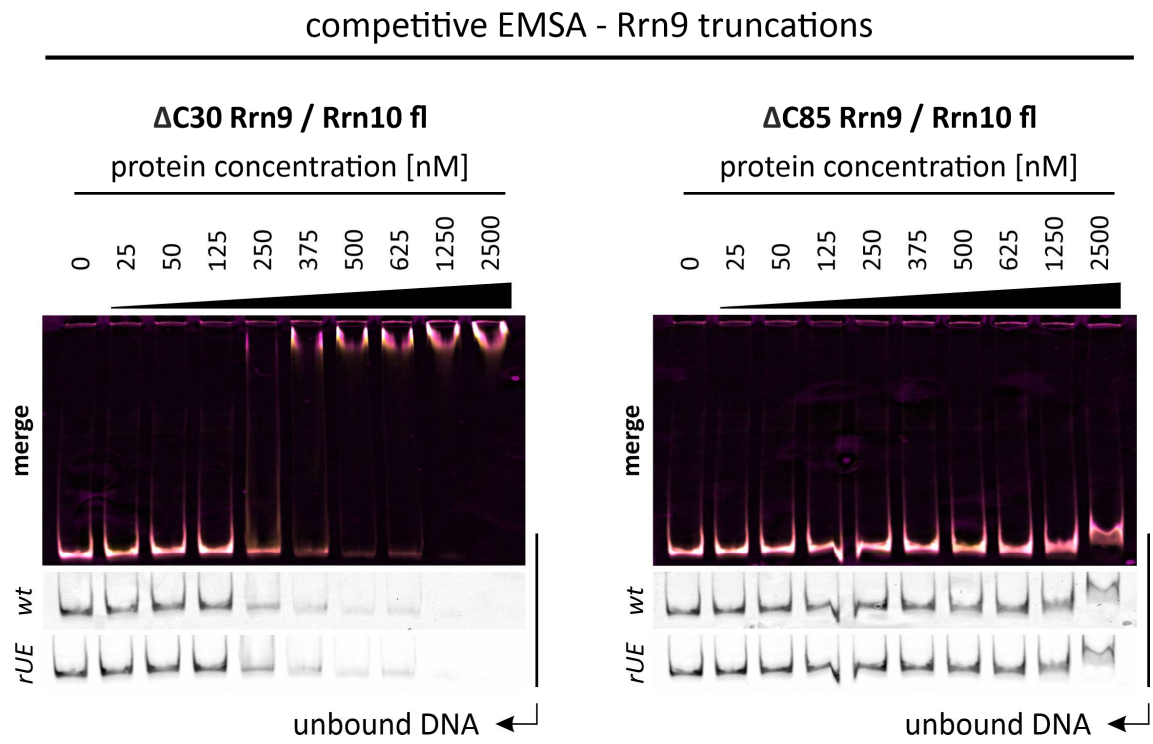


Figure 15: Competitive EMSAs of C terminal Rrn9 truncations.

Competitive EMSA of the Rrn9 Rrn10 subcomplex, run on a a 6% native polyacrylamid-gel. The wildtype (Cy3) and randomized upstream element (Cy5) template are represented in magenta and yellow, respectively (for reference see *Figure 13 A*. Signals were detected seperately (black/white channels) before they were merged in ImageJ. Absence of an upwards shift for the ΔC85 truncation indicates that the region between amino acid 280 (corresponds to the ΔC30 truncation – **left**) and amino acid 335 (corresponds to the ΔC85 truncation – **right**) is responsible for DNA binding. For construct comparison refer to Figure 7 and Supplementary Figure 4.

4.3.3 Uaf30 binds upstream promoter region with sequence specificity

DNA binding of Uaf30 was initially confirmed using standard EMSAs with the same templates (-155 to +8) as described in section *4.3.1*. Uaf30 showed binding towards both the wildtype sequence and the template containing the randomized upstream element in a similar concentration range from 250 – 625 nM (data not shown; refer to non-merged channels of competitive EMSA; *Figure 16*). However, in competitive EMSAs, in contrast to Rrn9/10, Uaf30 exhibited a slight preference for the wildtype sequence, evident from the subtle color change of the unbound DNA *(Figure 16)*. As with the Rrn9/10 subcomplex, we challenged this finding via 40 bp promoter fragment EMSAs and fluorescence anisotropy assays. Comparison of the unbound DNA in the promoterfragment EMSAs indicated a statistically significant preference for the DNA fragment spanning from -110 to -70 (F2) across 7 of the 11 tested concentrations *(Figure 17)*. T-tests were conducted for all pairs within each of the concentrations concentration. The corresponding p - values are provided in *Table 16*. Apart from the pairs listed in the table, no statistically significant differences in binding affinity could be detected.

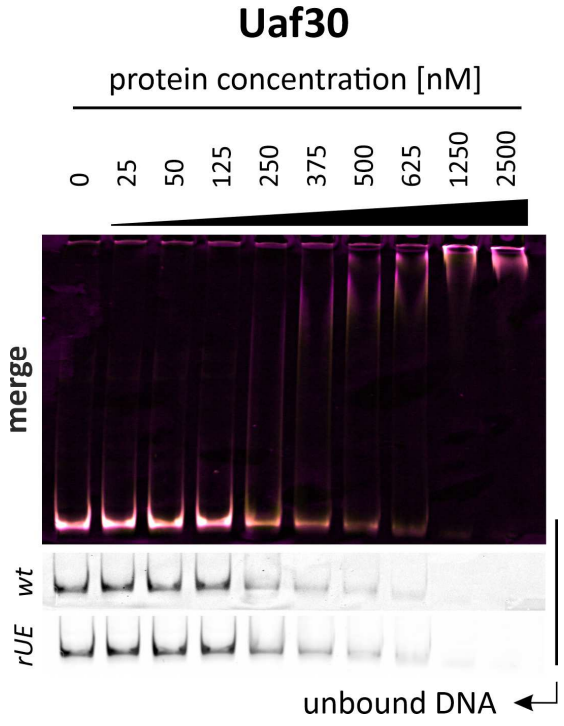


Figure 16: Competitive EMSA of Uaf30.

Competitive EMSA of Uaf30, run on a a 6% native polyacrylamid-gel. The wildtype (Cy3) and randomized upstream element (Cy5) template are represented in magenta and yellow, respectively (for reference see *Figure 13 A*). Signals were detected seperately (black/white channels) before they were merged in ImageJ. The faint color shift in the unbound DNA bands between 250 and 625 nM indicates a slight preference towards binding of the wildtype fragment.

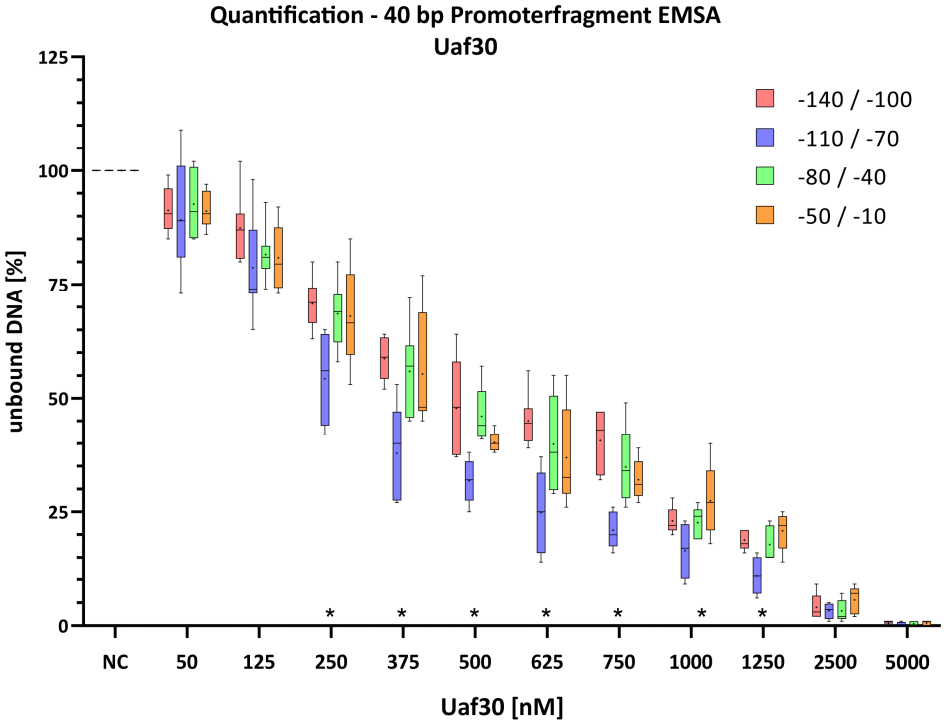


Figure 17: Quantification of Uaf30 promoterfragment EMSAs.

Unbound DNA of Uaf30 EMSAs for all promoter fragments was quantified and plotted for every concentration. EMSAs and quantification of unbound DNA were conducted as described in 3.6.1.2. Representative EMSAs are shown in *Supplementary Figure 8*. n ≥ 5 for all concentrations. Whiskers include all values (min to max), while the box extend from 25th to 75th percentile. The median value is indicated by a line inside the boxes, mean values by a +. Asteriscs above the x - axis mark all concentrations at which a statistically significant preference towards fragment #2 could be detected ($p < 0.025$). The respective p-values for each relevant tested pair are shown in *Table 16*. A clear preference could be detected for fragment #2, spanning the sequence from -110 to -70.

Statistical analysis of Uaf30 promoterfragment EMSAs

Table 16: p values for Uaf30 promoterfragment EMSAs.

Multiple comparison of means: Tukey contrasts; simultaneous tests for general linear hypothesis / T-tests were done for every pair inside every group (each protein concentration). * $p < 0.025$; ** $p < 0.01$; $p < 0.001$. The statistical analysis was done with help of Maximilian Pichler.

Conc.	Pair tested	p-value
250 mM	2 – 1	0.00236 **
	3 – 2	0.01001 *
	4 – 2	0.01383 *
375 mM	2 – 1	0.0043 **
	3 – 2	0.0127 *
	4 – 2	0.0154 *
500 mM	2 – 1	0.0106 *
	3 – 2	0.0243 *
625 mM	2 – 1	0.0032 **
	3 – 2	0.0298 *
750 mM	2 – 1	<0.001 ***
	3 – 2	0.0167 *
1000 mM	4 – 2	0.0145 *
1250 mM	2 – 1	0.00893 **
	3 – 2	0.02075 *
	4 – 2	0.00165 **

Fluorescence anisotropy measurements (n = 3) were additionally carried out to determine and compare the K_d values of all four assemblies, yielding results comparable to the promoter fragment EMSAs. Again, Uaf30 exhibited the strongest affinity towards fragment #2 (1379 ± 139 nM), followed by F4 and F3 with a decrease in binding affinity of approximately 600 nM each, respectively. With an affinity of ~ 5000 nM, the weakest binding was observed for fragment #4 *(Figure 18)*.

These findings collectively suggest that Uaf30, in contrast to Rrn9, binds the promoter DNA in a sequence specific manner, with a higher affinity towards an upstream promoter region between -110 and -70.

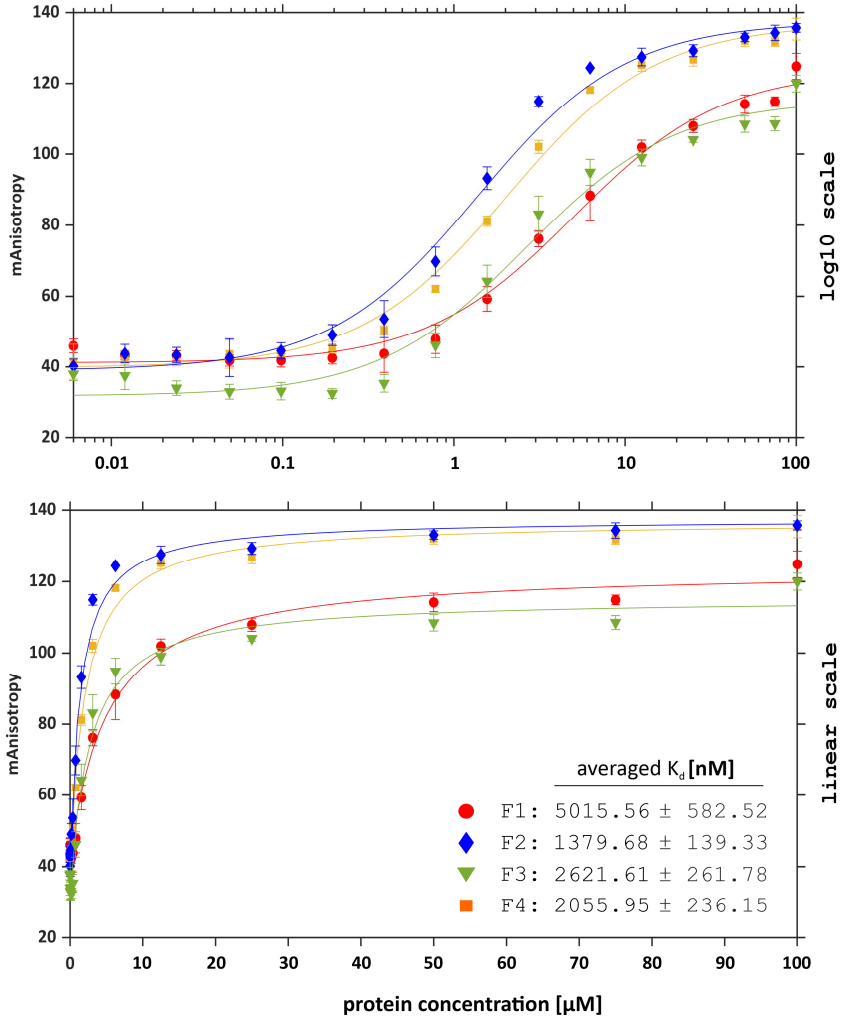


Figure 18: Promoterfragment fluorescence anisotropy of Uaf30.

Log10 and linear scale of fluorescence anisotropy measurements for K_d determination of Uaf30 with 40 bp fragments (for reference see *Figure 14 A)*. Technical triplicates of the anisotropy values were averaged (**m**eanAnisotropy – y-axis) and plotted against the protein concentrations. Curve fitting and the resulting K_d determination was done via MatLab with help of David Stelzig. (curve fitting script: 6.1.7). Note: protein concentrations (x - axis) are given in µM, K_d values in nM. Promoter fragment #2 (blue), exhibited the highest binding affinity towards Uaf30 (1379 nM), going in line with previous experiments *(Figure 17)*.

4.3.4 DNA binding by Uaf30 requires the full length protein

To determine which part of Uaf30 is responsible for DNA binding, EMSAs were conducted for each individual domain as well as for both domains incubated together. Interestingly, none of these approaches resulted in DNA binding *(Figure 19)*. From this, we conclude, that DNA binding of subunit Uaf30 likely requires the full length protein, including the putatively unstructured loop region between the N and C terminal domains *(Figure 26)*.

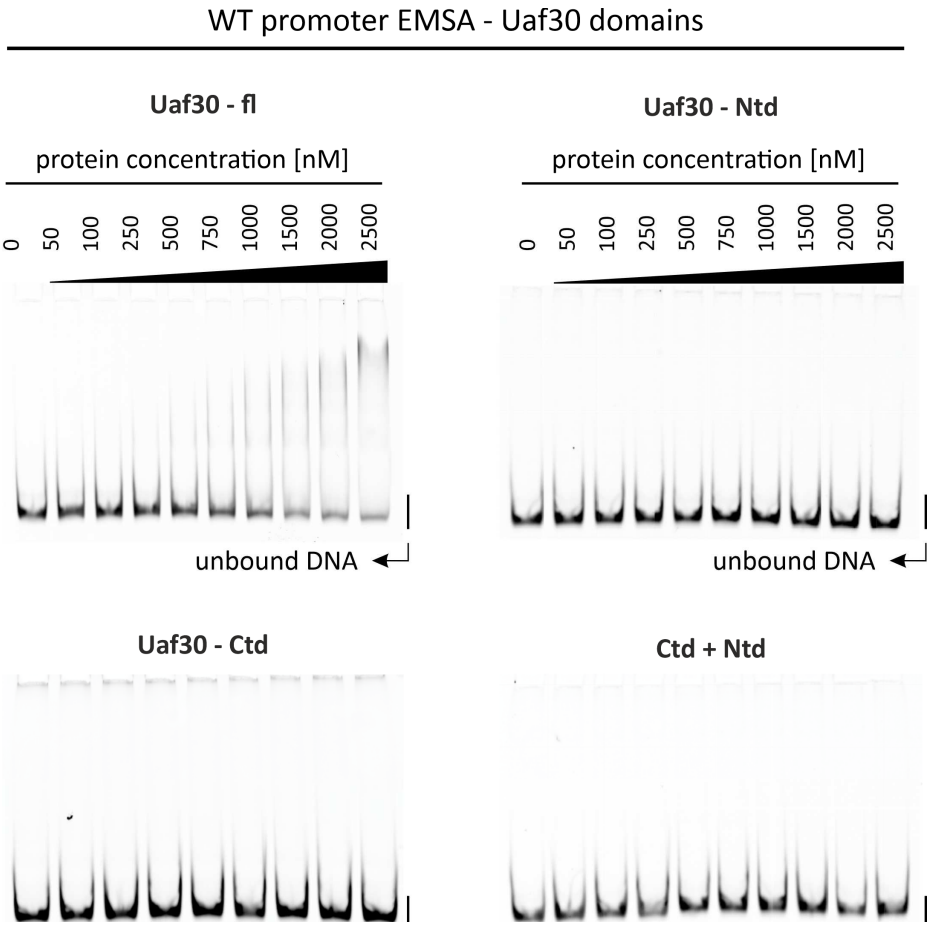


Figure 19: Wildtype promoter EMSAs of Uaf30 domains.

Standard EMSAs were done with a wildtype rDNA promoter template ranging from position -183 to +8 relative to the transcription start site with increasing amounts of Uaf30 and its N and C terminal domain respectively. Aside from testing the domains indivudually, they were pre – incubated prior to addition of the mastermix containing the DNA template. In contrast to full length Uaf30, none of the other approaches resulted in DNA binding, indicating that the full length protein is required for DNA binding.

4.3.5 UAF purified from *E. coli* specifically binds the upstream element

DNA binding and promoter specificity of the complete UAF complex were examined in competitive EMSAs. For this, UAF was incubated with two competing DNA templates of 331 bp length as described in *3.6.1.3*, with the wildtype upstream element sequence represented in magenta and the randomized UE probe in yellow (refer to template scheme in *Figure 14 A)*. As anticipated, a color transition from white to yellow could be observed for the unbound DNA *(Figure 20; left)*, clearly indicating a preferential binding of the wildtype sequence and verifying the functionality of the recombinant UAF complex, purified from *E. coli*. The unbound DNA in both channels was quantified for better visualization *(Figure 20; right)*. Additionally, ratios of unbound DNA template *(wt : rUE)* were calculated to improve comparability to the following competitive EMSAs of the deletion mutants (*secondary x – axis; lower values equal higher specificity*). Low values in the concentration range from 12.5 to 75 nM support the visual indications of the EMSA color transition.

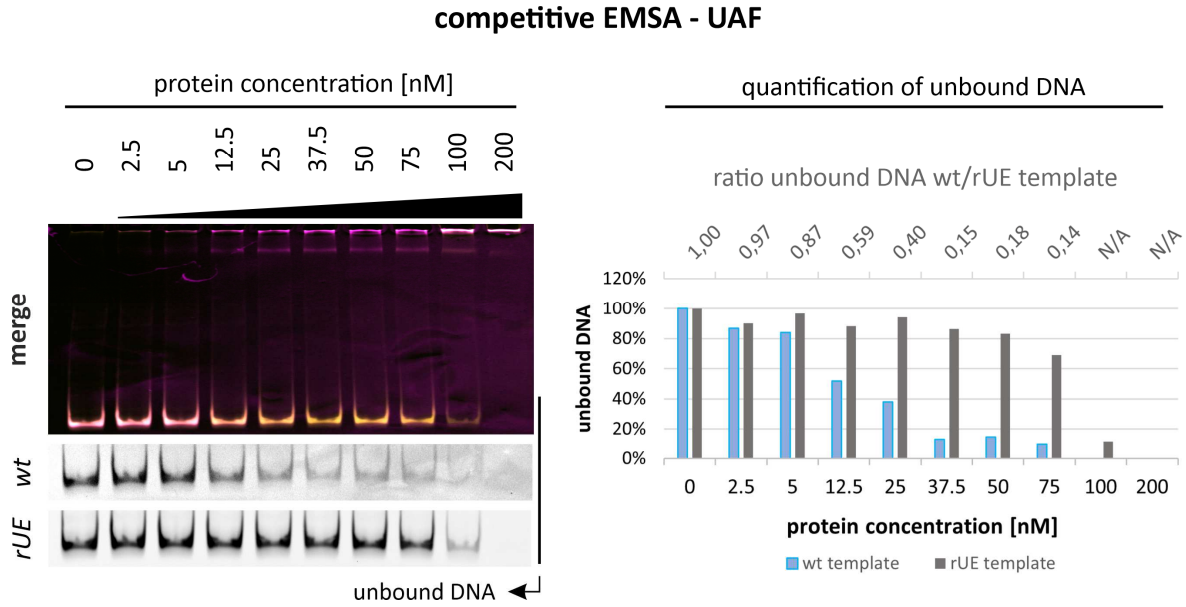


Figure 20: Competitive EMSA of whole UAF.

(left) Competitive EMSA of whole UAF, run on a a 6% native polyacrylamid-gel. The wildtype (Cy3) and randomized upstream element (Cy5) template are represented in magenta and yellow, respectively (for reference see *Figure 13 A)*. Signals were detected seperately (black/white channels) before they were merged in ImageJ. The color transition in the unbound DNA bands between 12.5 and 100 nM clearly indicates a strong preference towards the wildtype template. **(right)** Quantification of unbound DNA, seperately quantified for each channel (wt / rue). Values are normalized to the respective negative controls (lane 1) of both channels). Ratios of unbound DNA templates were calculated for better comparison of the binding behaviour. Low values indicate high specificity. The results show a strong binding of the wildtype template, as the control template is only bound as the wt template is almost completely occupied.

4.3.6 Deletions of Rrn9 – Ctd and Uaf30 impact DNA binding and specificity of UAF

With our analysis of the individual subunits we confirmed a specific binding of Uaf30 to a distinct 40 bp fragment inside the upstream element *(4.3.3)*. We were interested in further investigating the role of Uaf30 with a ΔUaf30 UAF mutant. Moreover, based on our observations in EMSAs using C terminally truncated Rrn9 *(4.3.2)*, we created a mutant lacking the C terminal region of Rrn9 that was responsible for DNA binding. Both mutants were also combined two a doublemutant. For this purpose, competitive EMSAs were conducted. To minimize pipetting and batch variations, the reactions of all complexes were done in paralell from the same reaction mastermix.

A deletion of the C terminal 85 amino acids of Rrn9 led to results, comparable to the full complex *(Figure 21 A + B, top)*. Full binding of the wildtype template was achieved at 75 nM, while the control template is fully bound at 200 nM. This would correspond to a higher affinity towards the wildtype template than exhibited by whole UAF (10% unbound DNA at 75 nM; full binding at 100 nM). However, small variations are not uncommon for this assay, especially as quantification of faint bands becomes increasingly imprecise. More importantly, throughout the lower concentrations of the ΔC85 Rrn9 deletion complex, a notable decrease of wildtype template binding was evident when compared to whole UAF *(Figure 20)*. This is also indicated by the higher ratios of unbound DNA templates (*secondary X axis*) in the concentration range from 5 to 50 nM. In line with our previous results *(4.3.1)*, the specificity of this complex towards the wildtype template was not reduced in this experimental setup, as the control template is only bound after the wildtype template is fully occupied.

Deletion of Uaf30 resulted in a pronounced decrease of specificity *(Figure 21 A + B, mid)*. The ratios of unbound templates demonstrate this effect particularly for concentrations ranging from 10 to 50 nM. At a concentration of 50 nM, for example, the ratio for the whole complex is 0.18, while the complex lacking Uaf30 exhibits a ratio of 0.89. Additionally, while the wildtype template was fully occupied at 100 nM for the whole complex, the Uaf30 deletion variant failed to fully bind the wildtype template until 200 nM. Most strikingly, in contrast to the ΔC85 Rrn9 deletion mutant, the complex lacking Uaf30 already interacts with the control template at lower concentrations, while the wildtype template is not yet fully bound, again pointing towards reduced specificity.

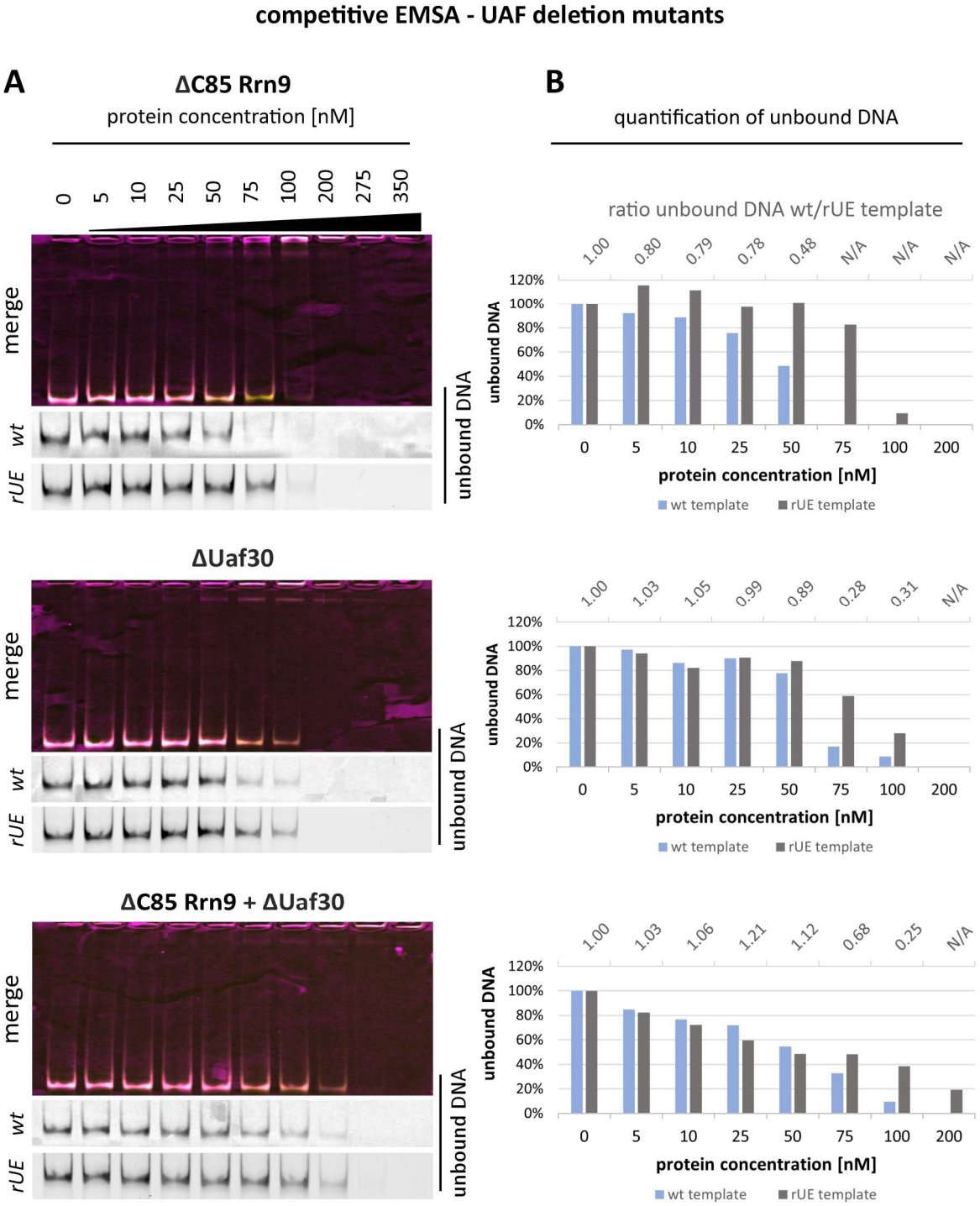


Figure 21: Competitive EMSAs of UAF deletion mutants.

(left) Competitive EMSAs of three UAF deletion mutants, run on a a 6% native polyacrylamid-gel. The wildtype (Cy3) and randomized upstream element (Cy5) template are represented in magenta and yellow, respectively (for reference see *Figure 13 A)*. Signals were detected seperately before they were merged in ImageJ. **(right)** Quantification of unbound DNA, seperately quantified for each channel (wt / rue). Values are normalized to the respective negative controls (lane 1) of both channels). Ratios of unbound DNA templates were calculated for better comparison of the binding behaviours. Low values indicate high specificity. The results indicate a notable decrease of specificity only for both variants lacking Uaf30, supporting previous findings *(4.3.3)*.

As expected, the combination of both deletions resulted in a complex that was significantly less specific to the wildtype template, as judged by the high ratios of unbound DNA in the concentration range from 5 to 75 nM *(Figure 21 A + B, bottom)*. Strangely, in this range, the overall affinity for both DNA templates was higher than the affinity of the ΔUaf30 complex. Yet, concentrations of 100 nM and more than 200 nM were required to fully bind the wildtype and control template respectively, indicating decreased overall affinity. Furthermore, ratios of unbound DNA for concentrations between 75 and 100 (200) nM indicate, that a certain degree of specificity towards the wildtype template seems to remain, even with both of our two identified DNA binding interfaces removed.

Taken together, these findings support our previous hypothesis about the DNA binding properties of Rrn9 - C and Uaf30.

4.3.7 Transcriptional activity is impaired in deletion mutants of UAF

To examine the impact that the two deletions may have on the transcription initiation rates by Pol I, in vitro transcription assays were performed as described in *3.6.3*. PIC assembly was was done in consecutive assembly steps, using two DNA templates of 331 bp (wt sequence) and 359 bp (randomized UE) length, according to the protocol established by Michael Pilsl as described in *3.6.3*. The reaction was started by addition of 32-P α-CTP for radioactive labeling of the two RNA transcripts (119 nt and 147 nt length respectively) *(Figure 22 A)*.

To first verify and compare transcriptional activity of the UAF complex that was purified from *E. coli*, it was compared to UAF purified from *Sf 9* insect cells, that was kindly provided by Michael Pilsl. Both complexes exhibited increasing transcriptional activity with higher protein concentrations (0.4 – 0.5 pmol per reaction), confirming functionality of the *E. coli* derived complex *(Figure 22 B)*. Negative controls included incomplete PIC assemblies lacking Net1 - C *(Figure 22 Lane #1)* CF *(#2)* TBP *(#3)* UAF *(#4)*. Furthermore, each DNA template was also separately transcribed with and without UAF for reference of the final transcripts *(Lane #9 – #12)*. The assay was repeated three times. Mean values of quantified signal intensities are shown in *Figure 22 D*.

The stabilizing and enhancing effect of Net1-C on Pol I transcription was extensively shown and discussed in Michael Pilsl's PhD thesis with the characterization of the rDNA promoter and other studies before (Hannig et al. 2019; Pilsl 2021; Shou et al. 2001). The effect also

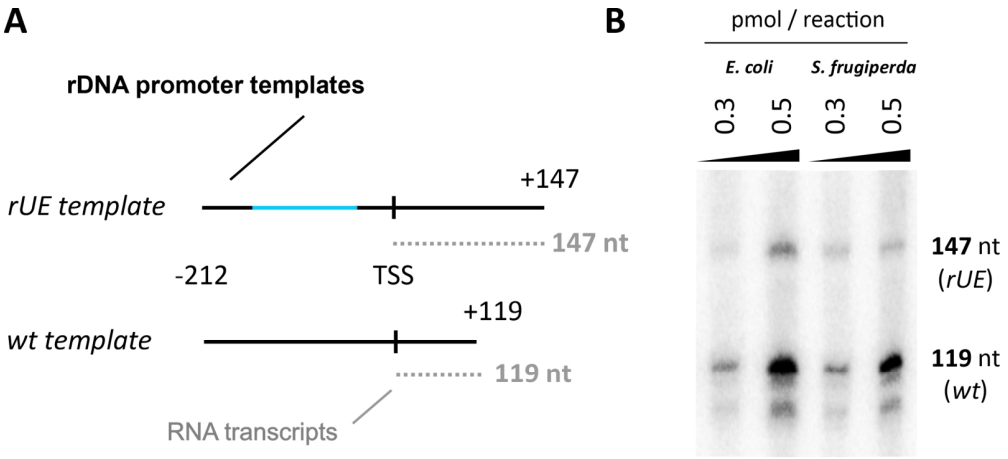
applies to the UAF complex purified from *E. coli*, which was evident from a comparison of Lane #1 (no Net1-C) vs. Lane #5 (+ Net1-C), both containing the whole set of transcription factors (TBP / CF / UAF / Rrn3). As anticipated, the control assembly lacking Core Factor was not able to transcribe either of the two templates *(Figure 22 Lane #2)*, and control reactions without TBP *(Lane #3)* or UAF *(Lane #4)* showed only basal transcription rates without template preference.

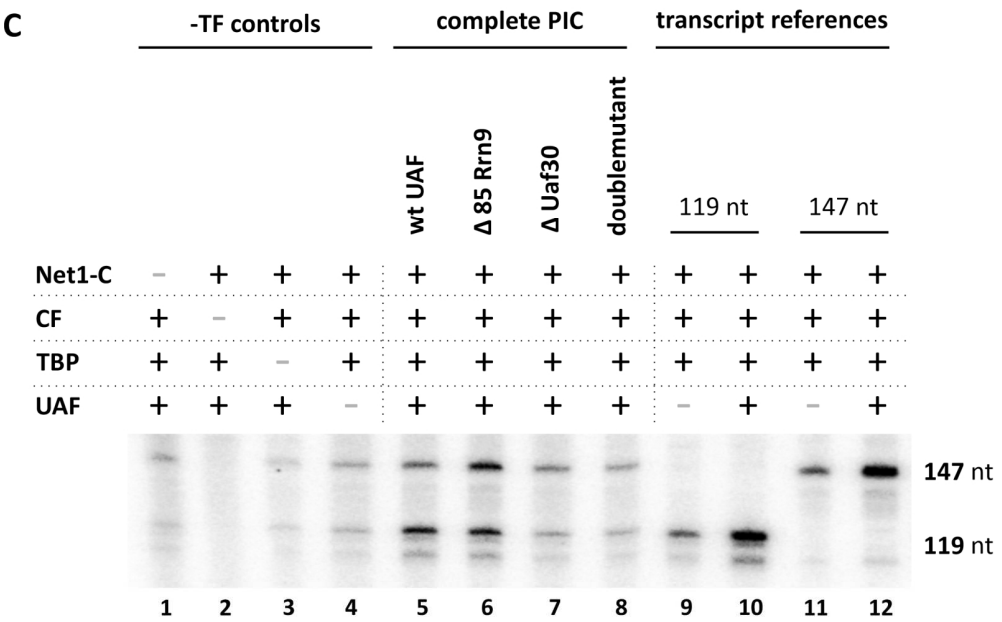
When adding whole UAF to the assembly *(Lane #5)*, a substantial increase (~ 5.5 fold) in transcription was observed (normalized to the 119 base transcript without UAF / *Lane #4*). Notably, although this increase was very specific to the wildtype transcript (119 nt), a modest increase of the control transcript (147 nt) was also observed (~ 1.8 fold) (normalized to the intensity of the control transcript without UAF / *Lane #4*).

Deletion of the C terminal end of Rrn9 *(Lane #6)* resulted in less enhanced transcription of the wildtype template, compared to wildtype UAF (~3.5 fold). Strikingly, the control transcript exhibited a similar increase in intensity (~3.4 fold). Deletion of Uaf30 from the complex led to a complete loss of transcriptional activity for the wildtype template *(Lane #7)*. A slight increase was however still observed for the control template (~1.4 fold). The doublemutant complex did not show any transcriptional stimulation, showing only basal levels of transcription with no preference towards either template *(Lane #8)*.

Taken together these results show, that in context of the whole Pol I PIC, UAF purified from *E. coli* can stimulate Pol I transcription with high specificity towards the wildtype promoter sequence in a competitive setup. Further, in line with the results of the DNA binding assays, the transcription enhancing effect of UAF, as well as its specificity are negatively impacted when one of the DNA binding regions is deleted.

However, these findings raise several questions that need further examination. The increase of intensity of the control template when adding both whole UAF and its ΔC85 Rrn9 variant is unexpected. This effect contradicts the results from our competitive EMSAs, which suggest that binding of the control template by this deletion complex does only occur when the wildtype template is fully bound. Furthermore, both single template controls exhibit the same amount of enhanced transcription when UAF is added. Given the lack of an upstream element in the control template, this effect is also rather surprising.


A
rDNA promoter templates
rUE template
+147
147 nt
-212
TSS
wt template
+119
119 nt
RNA transcripts
B
pmol / reaction
E. coli
S. frugiperda
0.3
0.5
0.3
0.5
147 nt (rUE)
119 nt (wt)


C
-TF controls
complete PIC
transcript references
wt UAF
Δ 85 Rrn9
Δ Uaf30
doublemutant
119 nt
147 nt
Net1-C
CF
TBP
UAF
147 nt
119 nt
1 2 3 4 5 6 7 8 9 10 11 12

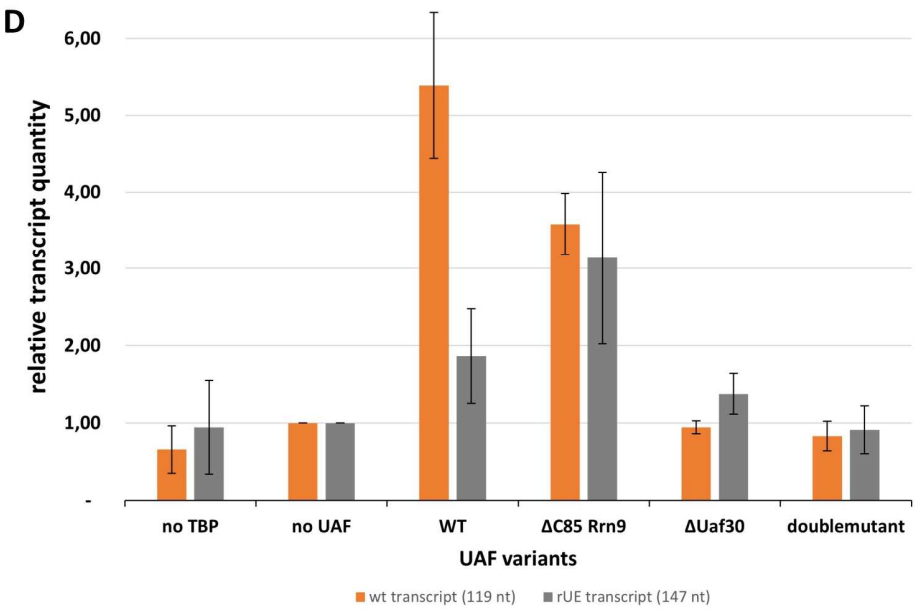

D
relative transcript quantity
6,00
5,00
4,00
3,00
2,00
1,00
-
no TBP
no UAF
WT
ΔC85 Rrn9
ΔUaf30
doublemutant
UAF variants
wt transcript (119 nt)
rUE transcript (147 nt)

Figure 22: In vitro transcription assay (Trx) of UAF and deletion mutants.

In vitro transcription of UAF and three deletion variants. **A)** Schematic representation of the DNA templates that were used for transcription. The randomized UE (randomized part: -155/-39 - blue) template results in a transcript of 147 nt length, the wildtype of 119 nt length. **B)** Direct comparison of UAF purified from *E. coli* and *Sf9* insect cells (M.Pilsl). Both complexes enhance transcription of the wildtype template, verifying functionality and comparability of the *E. coli* derived complex. **C**) Representative in vitro transcription assay: 32-P α-CTP radiosignal on an 8 % UREA – PAGE gel. Left to right: controls without each transcription factor (-TF) (lane 1-4); full PIC assemblies using four variants of UAF (lane 5-8); single template references without and with UAF respectively (lane 9-12). **D)** Quantification of relative signal intensities (n=3), normalized to PIC assemblies without UAF. Whole UAF (lane 5) notably enhances transcription of the wildtype template (~5.5 fold), while slightly stimulating the control template (~1.8 fold). Deletion of the Rrn9 DNA binding domain results in a complex that enhances transcription of both templates similarly (3.1 – 3.5 fold) (lane 6). Deletion of Uaf30, as well as the double deletion abolish the enhancing effect of UAF (lane 7).

4.4 Attempts for protein crystallization

In combination with a top down approach of analyzing UAF in context of the whole Pol I PIC via Cryo-EM (Michael Pilsl), a bottom up approach was pursued in order to gather structural data of UAF. For this purpose, we subjected all suitable purification products to high throughput crystallization screenings. The two main prerequisites were a minimum of ~95% purity and high solubility. An overview of all screens, proteins, and the tested concentrations is shown in *Table 17*. Generally, all proteins were first subjected to pre-crystallization tests (PCT), which helped to estimate a concentration range fit for each respective protein. Although most of the purified proteins had crystallization grade purity (SEC, SDS PAGE, western blots) and could be concentrated to a range that was suited for crystallization (according to the PCT), crystal formation could only be observed for a combination of C terminal truncations of the Rrn9/10 subcomplex: ΔC168 Rrn9 / ΔC25 Rrn10. In SDS PAGE, this combination showed no additional degradation bands *(Figure 9, gel 4)* and dimerization was significantly lower than for its longer counterpart *(Supplementary Figure 6)*. Four different conditions yielded a single crystal *(Figure 23)*. The respective negative controls (buffer + mother liquor) did not show any signs of crystal formation, which was a first indicator that the crystals did not form from buffer salts (di-Ammonium hydrogen phosphate (I) or $MgCl_2$ (III and IV)). Crystals were fished as described in 3.5.4 and flash frozen in different freezing buffers *(Figure 23)*. All crystals were analyzed on the Swiss Light Source (SLS) synchrotron at the Paul Scherrer institute in Villigen, Switzerland. None of the crystals that were anaylzed diffracted. In order to reproduce crystal growth and optimize the conditions

to yield diffracting crystals, HT screens were repeated using the same conditions. With this approach we were not able to reproduce crystal formation. Nevertheless, several grid screens were created to fine-screen around the initial hit conditions *(3.5.3)*. For the finescreening, drop size was increased to ensure pipetting consistency (0.3 + 0.3 µl to 0.5 + 0.5 µl) and hanging drop vapor diffusion was used instead of sitting drop.

Finescreens around initial hits shown in *Figure 23* covered: (I) $(NH_4)_2PO_4$ concentration from 500 – 1000 mM; NaAc concentration from 50 – 150 mM; pH from 4.0 – 5.5 (II) PEG200MME concentration ranging from 12.5 – 25 %, pH from 6.5 – 8.0 (III) and (IV) EtOH concentration ranging from 10 – 22.5 %, $MgCl_2$ concentration from 100 – 225 mM; pH from 7.0 – 8.5. Finescreenings were extended to (I – III) 15.5 and 18.5 mg/ml and (IV) 17 and 19 mg/ml, respectively. None of the tested conditions yielded crystals.

Table 17: Protein crystallization – High throughput (HT) screening overview.

List of all constructs, concentrations and commercial screens that were tested in the initial HT screening during this project. Conditions that yielded crystals *(see Figure 23)* are highlighted in green.

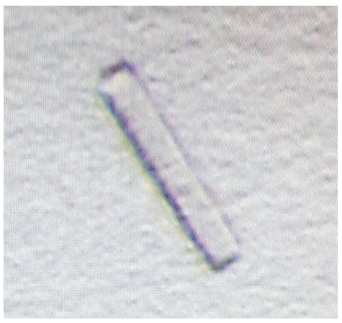
Protein	**[mg/ml]**	**Index**	**JCSG**	**PEG**	**$(NH_4)_2SO_4$**	**MPD**	**Wiz1+2**	**Wiz3+4**	**PACT**	**Morpheus**
Uaf30 – fl	5.6	•	•	•	•	•	•	•	•	•
	4.8	•	•	•	•	•	•	•		
	3.7	•	•	•	•	•	•	•		
Uaf30 – Ntd	1.4	•	•	•	•	•	•	•		
Uaf30 – Ctd	5.0	•	•	•	•	•	•	•		
	3.5	•	•	•	•	•	•	•		
Rrn9/Rrn10 – fl	21	•	•	•	•	•	•	•	•	•
	4.4	•	•	•	•	•	•	•	•	
	3.2	•	•	•	•	•				
ΔC30 Rrn9 / Rrn10 fl	17.5	•	•	•	•	•	•	•	•	
	8.75	•	•	•	•	•	•	•	•	
ΔC30 Rrn9 / ΔC25 Rrn10	16.7	•	•	•	•	•	•	•	•	
	8.5	•	•	•	•	•	•	•	•	
ΔC85 Rrn9 / Rrn10 fl	13	•	•	•	•	•	•	•	•	
	9	•	•	•	•	•	•	•	•	
ΔC85 Rrn9 / ΔC25 Rrn10	10	•	•	•	•	•	•	•	•	
ΔC168 Rrn9 / Rrn10 fl	11.2	•	•	•	•	•	•	•	•	•
ΔC168 Rrn9 / ΔC25 Rrn10	21	•	•	•	•	•	•	•	•	•
	17	•	•	•	•	•	•	•	•	•
	12	•	•	•	•	•	•	•	•	•
ΔN29 Rrn9 / Rrn10 fl	17.5	•	•	•	•	•	•	•	•	•
	13	•	•	•	•	•	•	•	•	•
ΔN29 ΔC30 Rrn9 / ΔC25 Rrn10	22.5	•	•	•	•	•	•	•	•	•
	17.5	•	•	•	•	•	•	•	•	•
ΔN29 ΔC168 Rrn9 / ΔC25 Rrn10	24	•	•	•	•	•	•	•	•	•
	18	•	•	•	•	•	•	•	•	•
C.t. Uaf30-like	8.7	•	•	•	•	•	•	•	•	
	5.8	•	•	•	•	•	•	•	•	
C.t. Rrn5-like Ntd	7.9	•	•	•	•	•	•	•	•	
	5.7	•	•	•	•	•				
UAF	5.5	•	•	•	•	•	•	•	•	•
	4.0	•	•	•	•	•	•	•	•	•

Crystals obtained from ΔC168 Rrn9 / ΔC25 Rrn10

17 mg/ml

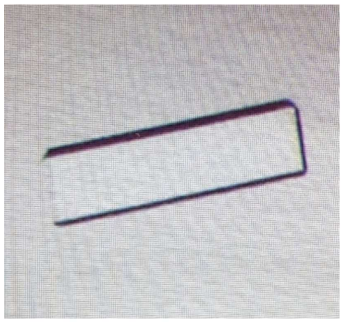
1 M di-Ammonium hydrogen Phosphate
100 mM Sodium acetate; pH 4.5

Cryo Buffer:
Screen buffer + 5% PEG 2000 MME



17 mg/ml

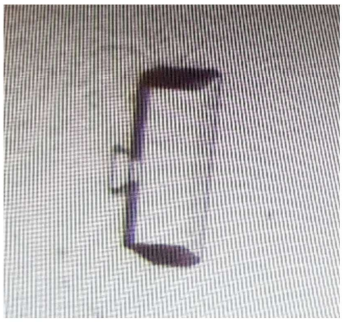
20 % w/v PEG 2,000 MME
100 mM TRIS; pH 7.0
Cryo Buffer:
Screen buffer + 10% glycerol



17 mg/ml

15 % v/v Ethanol
100 mM Imidazole; pH 8.0
200 mM Magnesium chloride

Cryo Buffer:
Screen buffer + 10% glycerol



21 mg/ml

15 % v/v Ethanol
100 mM HEPES; pH 7.5
200 mM Magnesium chloride

Cryo Buffer:
Screen buffer + 15% glycerol
+ 10% PEG 400

Figure 23: Crystals obtained from Rrn9/10 truncations.

Crystals that were obtained from the ΔC168 Rrn9 / ΔC25 Rrn10 subcomplex truncation in the initial high throughput (HT) screening at a concentration of 17 and 21 mg/ml. Screen conditions for each hit are listed under each respective picture. All crystals were only found once per condition. Different cryo buffers were tested for freezing and storage in a liquid nitrogen tank until they were analyzed at the synchrotron.

5 Discussion

5.1 Rrn9 / Rrn10 purification issues

Structural analysis via crystallography requires large amounts of highly concentrated and pure protein. For this reason, endogenous expression in the native system of UAF (*S. cerevisiae*) was unfit for a crystallographic approach. We established and optimized protocols for overexpression and purification of the recombinant Rrn9/10 subcomplex and Uaf30 from *E. coli*. Furthermore, protocols were created for an N (Uaf30 – Ntd) and C (Uaf30 – Ctd) terminal domain. This system also allowed for uncomplicated construct engeneering for functional analysis.

Purification of the Rrn9/10 subcomplex always yielded a product that was stably interacting but showed C terminal degradation of both subunits. This has been observed by other research groups before (Knutson et al. 2020; I. Siddiqi et al. 2001; M. L. Smith et al. 2018), but poses issues that have to be discussed. Firstly, any form of heterogeneity inherently diminishes the likelihood of successful crystal formation. This issue was sucessfully adressed by truncating both subunits, resulting in more stable subcomplexes, one of which eventually yielded crystals *(Figure 23)*. More importantly, one must consider, that results that are obtained from functional assays using this heterogenous Rrn9 / Rrn10 subcomplex will always remain questionable in regards to potential impacts of this heterogeneity.

As described in section 4.2.1, an additional challenge was the observed time dependent multimerization of the Rrn9 / Rrn10 subcomplex *(Supplementary Figure 6)*. Immediate freezing of monomeric Rrn9/10 after a second size exclusion run helped to avoid this issue, at least temporarily. Functional assays with Rrn9 / Rrn10 were done as quickly as possible after thawing aliquots of purified Rrn9/10 to minimize result distortion by this effect. According to the retention volume, the second highest peak would correspond to formation of dimers. Given the structural similarity of Rrn9 and Rrn10 to histones H2B and H2A respectively (Baudin et al. 2022) *(Figure 4 D)*, it is not unlikely that this dimerization resembles histone assembly or assembly of the core complex of UAF as described in the introduction *(2.4.1)*. Disruption of the histone fold domain of Rrn9 in the shortest construct (for reference see *Figure 7)* showed significantly reduced rates of dimerization *(Supplementary Figure 6)*, which would go in line with this hypothesis. However, other effects

that could promote dimerization, e.g. newly formed unspecific protein - protein interactions of the subcomplex with itself, can not be ruled out.

As mentioned above, truncating the C terminal end of Rrn9 and Rrn10 resulted in more degradation resistant complexes, while maintaining their stable interaction *(Figure 9)*. Previously, it was reported that this interaction is strongly impaired (yeast 2 hybrid assays) when residues L185 and F186 or Rrn9 are mutated to serine (Steffan et al. 1998). Our results do not support these findings, as a C terminal truncation up to residue 172 did not result in an evident stoichiometric imbalance of the subcomplex in an SDS PAGE after purification *(Figure 9, gel 2)*. Recent crosslinking studies also showed, that Rrn10 association to Rrn9 was only reduced for a Rrn9 mutant lacking the region between 71 – 111 (Knutson et al. 2020), going well in line with the slightly larger interface we could identify with our Rrn9 truncations (29 – 172). However, as yeast 2 hybrid assays represent the in vivo situation, direct comparison with assays using purified components in vitro is difficult. Various secondary and/or tertiary factors might influence UAF assembly e.g. due to incorrect folding of Rrn9.

5.2 Functional analysis of UAF subunits

5.2.1 Rrn9 – C and Uaf30 are required for promoter binding

rDNA promoter binding of the recombinant Rrn9/10 subcomplex was examined via different EMSAs. First, in addition to the wildtype promoter template, a control template was tested. This second template contained a randomized sequence from -155 to -38 (sequence alignment: *Supplementary Figure 7*). This control template was regularly used by Michael Pilsl in order to examine the binding behavior of UAF towards the promoter vs. the control sequence (Pilsl 2021). Other studies used truncated templates that lacked the region upstream of certain positions in the promoter (Keener et al. 1998; Kulkens et al. 1991; Musters et al. 1989), leaving more room for interpretation regarding the role of potential cooperative unspecific interactions in this deleted upstream region. We could not detect specificity of the Rrn9/10 subcomplex towards the wildtype template in competitive EMSAs, nor could we see any preference for any of the 40 bp fragments spanning the upstream and core element of the wildtype promoter sequence. From these results, we concluded that the Rrn9/10 subcomplex binds the promoter DNA in a sequence unspecific manner. MST and fluorescence anisotropy measurements were attempted to further support the EMSA data.

In both cases the required concentrations for proper determination of the binding affinity greatly exceeded the amounts that were needed for binding in EMSAs, rendering yield and concentration of the protein batches the limiting factors for a reasonable experimental setup.

Stepwise C terminal truncation of Rrn9 showed, that its binding capability relies on a region between amino acids 280 and 335, as introduction of a stop codon at position 280 resulted in a loss of DNA binding, while a stop codon at position 335 did not. This finding goes well in line with the recently published model that reports a binding interface between Rrn9 and promoter DNA, which relies on this exact region of Rrn9 *(Figure 24; C terminal deletion colored in gray)* (Baudin et al. 2022). However, this model also suggests two interactions between an N terminal helix (specifically: amino acid Y89 and K86) and the phosphate backbone at position -55 and -54 respectively (*Figure 24; green circle*). Although a loss of binding was observed for deletion of the C terminus, we cannot exclude, that the N terminal helix plays a role in promoter binding. Higher concentrations (> 2.5 µM in EMSAs) might potentially result in binding of the N terminal helix, which could exhibit a lower affinity that was overlooked in our setup. One could also argue that the deletion may lead to conformational changes in the adjacent elements, e.g. the N terminal helix, resulting in an altered structure of the subunit, no longer able to bind DNA. The atomic model also shows one base specific interaction at position position -57 (K308 of Rrn9). Although our results suggest nonspecific binding, we cannot rule out the possibility that one specific interaction may not suffice to produce a distinguishable specific binding event, while primarily relying on unspecific phosphate backbone interactions (ratio 1 : 5). Also, as stated above, structural changes caused by the deletion are conceivable. Interestingly, promoter mutant analyses done by Michael Pilsl have shown slightly reduced specificity of UAF (compared to the wildtype sequence) when substituting the region between -68 to -39 with a randomized sequence (Pilsl 2021). However, this effect is not necessarily applicable to Rrn9 individually, which might explain the lack of specificity in this experimental setup. The role of Rrn9 in context of whole UAF will be discussed in the next section.

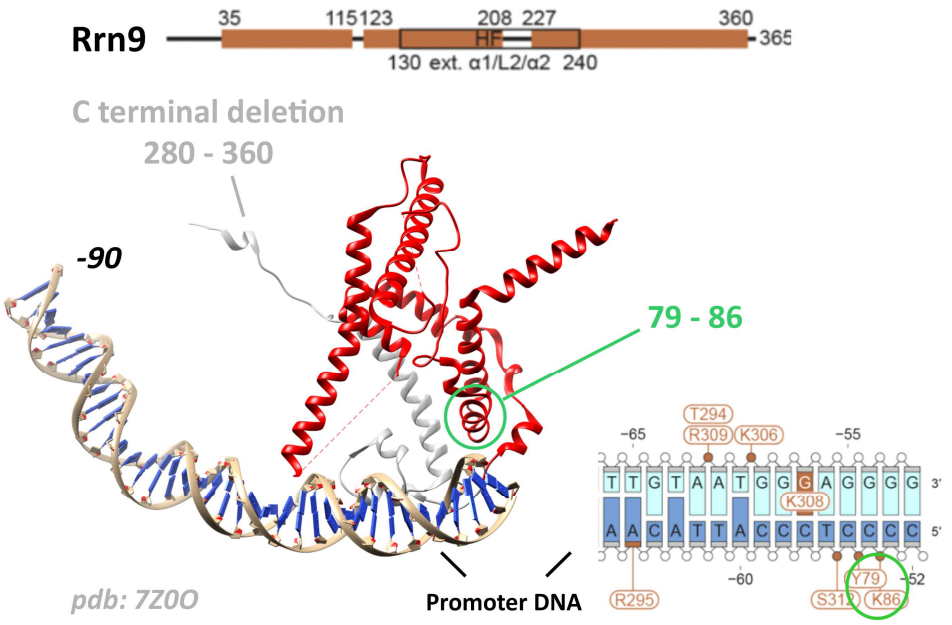


Figure 24: Review of Rrn9 rDNA promoter contacts.

Visualization of the deletion of the DNA binding domain of Rrn9 using a modified pdb model of UAF and the promoter DNA from Baudin et al. as shown in *Figure 4*. The C terminal end of Rrn9, that is deleted in the truncation that loses its DNA binding ability, (see chapter *4.3.2*) is colored in light gray. N terminal amino acids putatively forming phosphate backbone contacts are highlighted with a green circle. Bottom right: Protein DNA contact positions as stated in the respective publication. N terminal contacts are highlighted as in A. - Baudin et al. 2022

Functional analysis of Uaf30 via different binding assays revelaed a specific protein DNA interaction with a preference towards an upstream region between -110 to -70. Notably, as already mentioned for Rrn9/10, the concentration that was required for fluorescence anisotropy (FA) measurements and reasonable curve fitting was 10 – 20 x larger than for the EMSAs. However, in contrast to Rrn9/10, a full binding plateau of the promoterfragments could be achieved with this subunit. Hence, fitting of a sigmoidal curve and K_d determination were possible. Although the highest affinity can be attributed to the promoter fragment from -110 to -70 in promoter fragment EMSAs as well as fluorescence anisotropy, affinities of the other three fragments are relatively different in FA *(Figure 18)*. Possibly, secondary factors also play a role in the DNA binding of Uaf30, at least in this particular assay. One could speculate that plasticity of - or secondary structures in - the template DNA influence how efficiently the subunit can bind DNA. Thus, the varying affinities between F1, F3, and F4 (and also F2) may still be caused by sequence dependend differences in DNA shape (e.g. hairpin structures, loops, bendability). It should however also be noted, that, regarding secondary

structure formation, 40 bp fragments likely behave differently than the full length promoter, or longer DNA templates in general.

In the newly available structure of UAF, due to high flexibility of the central and C terminal region of Uaf30, only the N terminus is resolved (amino acid 1 – 83) (Baudin et al. 2022). However, the supplementary data shows the electron density of a subpopulation of particles in the region between the N terminus of Uaf30 and upstream promoter DNA around the positions -100 and -90 *(Figure 25)*. It is speculated that this density is occupied by the flexible C terminus of Uaf30. This would go in line with our observations, as this region is covered by promoter fragment #2 (-110 to -70), showing the highest affinity binding for Uaf30. DNase footprinting assays by another group, that have already suggested a binding site in this area, revealed an uncovered region when comparing whole UAF to a ΔUaf30 mutant (hypersensitive site at -107/-97) (Hontz et al. 2008). It was however unclear, if in context of the whole UAF, this region is exposed to DNase due to structural alterations in the mutated complex or because a direct contact between Uaf30 and DNA was absent. Here, we could show a direct interaction between Uaf30 and the promoter DNA and confirm a higher affinity binding site in this region, which requires the full length protein for successful assembly.

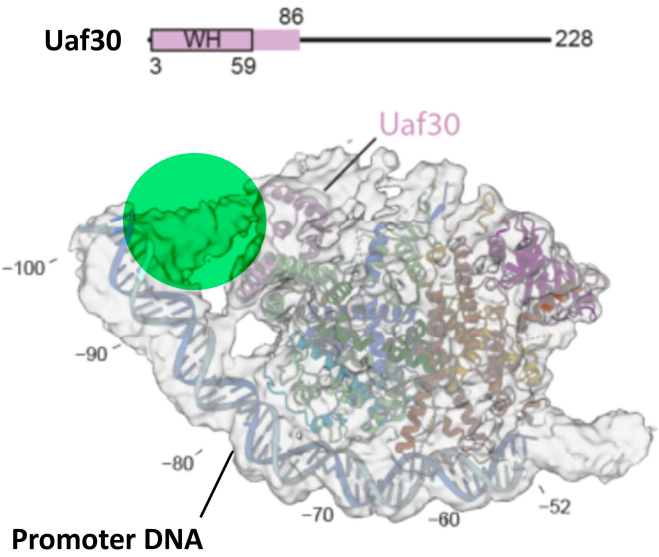


Figure 25: UAF electron density map showing putative C terminal domain of Uaf30.

Electron density map with fitted UAF of a subpopulation of particles detected during Cryo-EM reconstruction of UAF/TBP and promoter DNA shows low resolution density in viscinity of the N terminus of Uaf30. It is speculated that this density represents the flexible C terminal end of Uaf30. (adapted from Baudin et al. 2022)

5.2.2 Deletion of DNA binding regions affect affinity & specificity of UAF

In previous experiments done by Michael Pilsl (MP) with whole UAF purified from insect cells, randomizing the upstream element (UE) resulted in a ~10 fold reduced binding affinity. Competitive EMSAs with MPs complex showed binding of the randomized UE template only after the wildtype template was almost completely bound (Pilsl 2021). We were able to reproduced this result with the UAF we purified from *E. coli (Figure 20)*.

For direct comparison, the same assay was simultaneously performed for all of the three mutant complexes, also purified from *E. coli*. In agreement with the results for subunits Rrn9/Rrn10 and Uaf30, the mutant complexes were affected in regard to their binding affinity and specificity. These results strengthen our hypothesis that Uaf30 is providing specificity towards the promoter, while Rrn9 binds cooperatively but unspecifically.

Although, in the double mutant, two important DNA binding interfaces of UAF were disrupted, it was expected that (unspecific) binding would not be completely abolished. However, it is interesting to see that UAF maintains a slight preference towards the wildtype template. The UAF model from 2022 reports several additional DNA contacting residues (Rrn5 and H3) in the region from -85 to -68 nucleotides upstream of the TSS. Further, as discussed above, an N terminal helix of Rrn9 is also reported to engage in contacts with the promoter (Baudin et al. 2022). Although all of these are reported to be phosphate backbone or sugar interactions, and would hence be sequence unspecific, our data indicates that this remaining sequence specificity might nevertheless have its roots in one of these remaining UAF - DNA interfaces. However, since our experiments do not cover these interactions, we cannot attribute this remaining specificity to either of these elements without uncertainty. Aside from that, as previously mentioned, sequence dependent plasticity of the DNA or formation of secondary structure elements are neglected as a factor in our experimental setup.

To conclude, the real nature of the remaining specificity towards the wildtype sequence can not be answered with our findings. Followup assays could narrow down the origin of this specificity. A deletion of the DNA contacting N terminal helix of Rrn9 could, for example, help to clarify the role of the helix in sequence specificity of UAF. Likewise mutation of Rrn5 and H3 residues that are reported to contact the DNA, could do the same. Furthermore,

randomizing the DNA stretch between -110 / -70 could be a good control to verify the specificity of Uaf30.

In fact, altering or deleting stretches inside the promoter sequence has been done before. However, the existing data is not unambiguous. Deletion of the 5' region from -208 to -91 was reported to result in a strong decrease of template activity in template competition assays versus the wildtype control template. This decrease was not observed when only deleting the region upstream of -155, suggesting that the region between -155 and -91 plays a crucial role in promoter targeting and specificity (D. A. Keys et al. 1996). These results go in line with our findings regarding the role of the promoter region around position -100. Transcription assays that were done using purified components of the Pol I PIC contradict these results. Here, 5' truncations of the promoter up to -76 barely affected transcription stimulation by UAF, compared to a template ranging up to -210 (Keener et al. 1998). Truncations reaching even further downstream, up to position -60 and -38 (the latter corresponds to a deletion of the whole upstream element), almost completely abolished stimulation by UAF.

Interestingly, cis element analysis from our lab, carried out by Michael Pilsl, showed a rather strong decrease (> 50 %) in wildtype template specificity for deletions of the upstream region between -155 to -76. The wildtype template was however still preferentially bound by UAF (Pilsl 2021). As already mentioned in the previous chapter, a deletion from -68 to -39 also reduced sequence specificity by UAF. While this reduction was not as prominent as for the deletion of the upstream region around position -100, it might nevertheless be an indicator of a slight sequence specificity of Rrn9. As already pointed out before, in the atomic model, this promoter region is associated with Rrn9 interactions. It should be noted, that, in contrast to the studies mentioned before, in this case a randomized sequence was substituted for the deleted parts of the promoter to provide templates of the same lengths.

To sum up, although conflicting data exists on the exact sequence stretches that are important for UAF binding and specificity, the majority, including our new results, show that the region around position -100 plays an important role for promoter recognition by UAF. The results presented in this thesis are not unambiguous in regard to the degree to which Rrn9, Rrn5, and H3 contribute to sequence specificity. Although we cannot detect specific binding in EMSAs or promoter fragment EMSAs for the Rrn9/10 subcomplex, this can not necessarily be applied to the the function of Rrn9 in the whole complex, as structural

differences cannot be excluded. Since we still observe specificity when Uaf30 and/or the C terminal region of Rrn9 are deleted, and our experimental setup does not allow conclusions about the source of this specificity, we can only speculate that Rrn5, H3 or the N terminal helix of Rrn9 are responsible. To verify this hypothesis, further experiments have to be done.

5.2.3 Transcription stimulation by UAF and deletion variants

The augmenting effect that UAF has on Pol I transcription has been examined via in vitro transcription assays (Trx assays) by several groups in the past, including our own (Michael Pilsl). Most of the other studies centered around the whole UAF complex and primarily focused on testing rDNA templates, omitting specific regions of the promoter through truncations (Keener et al. 1998; Kulkens et al. 1991; Musters et al. 1989). In contrast to their approach, and similar to the approach of Keys et al. (1996) our experimental setup involved two template strands, one with a randomized upstream element (UE) (Pilsl 2021). With the aim of understanding the origin of the specificity of UAF, the sequence randomization approach seems like a more elegant solution than a full truncation of the template. However, given that a small percentage of the randomized sequence contains bases identical to the wildtype UE, a small chance remains that single base specific interactions still occur with the control template (sequence alignment *Supplementary Figure 7)*. This factor could be eliminated by designing a fully non-identical control template.

In our lab, the Trx protocol was established by Michael Pilsl (MP) who used UAF from insect cells to characterize cis elements inside the rDNA promoter during his PhD (Pilsl 2021). We compared the transcription enhancing effect of both complexes (insect & bacterial), verifying the functionality of the *E. coli* derived UAF. As described in the introduction, the enhancing effect of UAF was reported to reach levels of up to a 10 to 50 fold increase (Keener et al. 1998), compared to a 5.5 fold increase that we could observe. However, in their setup, endogenous UAF was purified from yeast. Furthermore, stoichiometries that were used in our Trx assays were titrated using proteins from our lab own purification protocols that were established by Michael Pilsl (Net1-C, CF, TBP, Rrn3, Pol I). Given the complexity of the assembly of the complete Pol I PIC, small deviations from the original protocols that were used in the previously mentioned publication may influence transcriptional efficiency. Aside from this, efficiency of PIC assembly and its functionality could also suffer from suboptimal

preconditions, such as varying quality and unfit stoichiometry of each transcription factor, or simply freezing damage.

In addition to the expected enhancement of the wildtype template transcription, the bacterial complex exhibited a notable enhancing effect on the control template. Furthermore, when testing both DNA templates individually, addition of UAF had the same enhancing effect on both templates. This was unexpected, as the control template lacks the wildtype sequence of the upstream element, and no such effect was observed for the insect UAF, used in MPs studies (Pilsl 2021). To identify the cause for this effect, follow up experiments have to be done. Approaches will be discussed below.

With the knowledge we gained from the characterization of Rrn9, we wanted to extend this in vitro Trx approach to our newly mapped DNA binding domain in the C terminal region of Rrn9. As discussed above, the wildtype complex unexpectedly enhanced transcription of the control template. This effect was considerably more evident for the ΔC85 Rrn9 variant. Interestingly, while the intensity of the wildtype transcript still remains slightly higher, the transcript ratio was almost 50:50, indicating that the deletion strongly reduced the specificity of the complex. As no Trx data exists for any kind of Rrn9 deletion complex, this result in itself is intriguing. However, it contradicts our previous results from the characterization of Rrn9/10 individually and in context of the whole complex, which collectively suggest, that the specificity of UAF is likely not mediated by Rrn9. Especially the competitive EMSA, which showed that binding of the control template starts after complete binding of the wildtype template. The reason for this remains unclear. Possible follow up experiments for further investigation of this effect will be discussed below.

The idea of deleting subunits of UAF and study the effect it has on the transcriptional activity of Pol I is not novel (Siddiqi et al. 2001; Smith et al. 2018). For example, in a publication by *Siddiqi et al*., transcriptional activity of a complex lacking Uaf30 was compared to the full complex. They demonstrated, that the deletion of Uaf30 from the complex leads to a substantial reduction (- 50 %) in transcriptional enhancement. However, in contrast to the ΔUaf30 complex that was used in this work, their complex was purified endogenously from yeast. Still, the near - complete loss of enhancement in our experimental setup, using UAF purified from *E. coli*, raises the question of the source of that striking discrepancy. As mentioned in the introduction, possible posttranslational modifications are not taken into account in this setup. For example, histone tail modifications (e.g. acetylation) have been

shown to affect protein and DNA interactions (Allfrey et al. 1964; Eberharter and Becker 2002; Luger et al. 1997). It is conceivable, that, given the presence of 3 histone subunits in the complex, the deletion of Uaf30 leads to a loss of function that is caused by structural alterations that do not occur when deleting the subunit in an eucaryotic expression system. If this is the case, structural analysis of the ΔUaf30 complex from *E. coli* is necessary to identify changes in the overall structure of the complex. Further, more information about posttranslational modifications of UAF could help to understand the lack of transcriptional enhancement in our assays.

Taken together, these Trx results raise several questions, that need to be adressed in future experiments:

1. Why does the deletion of Uaf30 abolish the transcription enhancing effect of UAF in this setup, using a bacterial expression system?
2. Why does the deletion of the DNA binding domain of Rrn9 lead to decreased specificity, although we could not see specific binding in other assays?
3. Why does UAF from *E. coli* enhance transcription of the control template, while UAF from insect cells does not?

The following approaches could help to answer these questions:

1. As stated above, expression of a ΔUaf30 complex in the established baculo system could help to clarify if the loss of the enhancing effect of this mutant complex is caused by expression in a procaryotic system, possibly due to missing posttranslational modifications. Furthermore, obtaining structural data of ΔUaf30 complexes from both expression systems could help to examine structural alterations of the residual subunits when Uaf30 is deleted from each complex.

2. Conducting competitive EMSAs in parallel to the in Trx assays, taking a small fraction of each reaction, would be a suitable control to monitor template occupation and verify, that the starting point matches the situation from the competitive EMSAs that were done seperately. Furthermore, in order to comprehend possible structural changes in the complex that are caused by a deletion of the C terminal region of Rrn9, the structure of a complex carrying this deletion needs to be resolved.

3. Screening of different DNA templates could help to improve our understanding of the ability of UAF to form sequence unspecific interactions with DNA in absence of a UE, followed by a full assembly of the PIC. A truncated template starting at position -38, lacking the complete upstream DNA, would not be able to form such a complete PIC. Furthermore, as already suggested, a complete randomization would eliminate the residual risk of sequence specific interactions between UAF and the UE control template. A second control template, harboring an inversed upstream element could lead to a reversed orientation of UAF, and thus supression of its transcription enhancing effect.

5.3 Problems with - and possible solutions for – crystallization

All purified proteins were first analyzed in pre crystallization tests before they were subjected to various high throughput (HT) buffer screens (Watson and O'Callaghan 2005) (Herrmann and Bucksch 2014). Approximating a starting concentration through PCT helped avoiding a serious overshoot or shortfall in protein concentration for the HT screens. However, this was only a rather vague estimation and left a wide range of possible concentrations to start with.

In addition, for the Rrn9 / Rrn10 subcomplex, the influence of both issues that were discussed above were incalculable. Although dimerization itself is not necessarily bad for crystal formation, as dimers can also form self repetitive patterns, the combination with the strong C terminal degradation of both subunits renders formation of a homogenous repetitive crystal lattice unlikely.

However, a fusion of the ΔC168 Rrn9 and the ΔC25 truncation of Rrn10 yielded a subcomplex that was substantially decelerated in dimerization and showed no signs of degradation in SDS PAGE *(Supplementary Figure 6, Figure 9)*. This truncated subcomplex yielded crystals under different buffer conditions *(Figure 23)*. Unfortunately, the exact cause behind the lack of crystal diffraction and the inability to reproduce or optimize these crystals remains unclear.

Different approaches could have helped to further optimize protein stability prior to the high throughput screening or in drop after setting up plates. It has been shown that reductive surface lysine methylation can increase the probability of crystallization by providing new contacting points at lysine clusters (Sledz et al. 2010) (Walter et al. 2006). A major drawback

of this method is the need for even more protein, as it is incrementally lost with an increasing number of preparative steps. Nevertheless, this approach could be worth a try, especially because Uaf30, the subunit missing the largest fragment in the published atomic model, is predicted to be a suitable candidate for this protocol. A freely available online tool (surface entropy reduction prediction server "SERp" (Goldschmidt et al. 2007) predicts four lysine clusters, mostly in disordered regions of Uaf30, that could be methylated *(Figure 26 A + B)*.

Although the creation of new surface contacts might lead to a protein structure that deviates from the native positioning of both domains in relation to each other, the resolution of the N and/or C terminal domain could potentially be achieved through this approach, if crystallization is thereby facilitated.

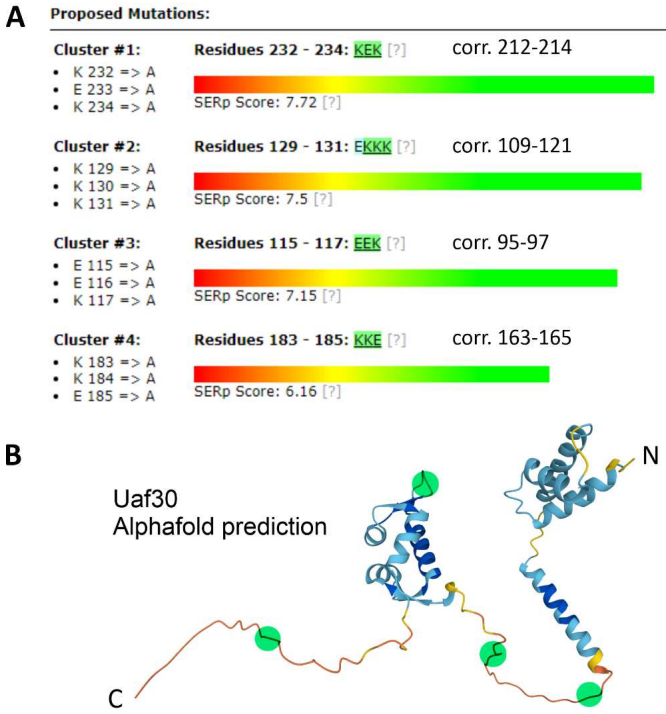


Figure 26: Surface entropy reduction (SERp) lysine cluster predictions for Uaf30.

A) Uaf30 lysine clusters that were predicted using the SERp online prediction tool. The position of each cluster is shifted 20 amino acid positions due to the included his tag in the fed sequence. Corrected positions are stated on the right. The location of each cluster is highlighted in B. **B)** Alphafold prediction of Uaf30 (dated 03/2023) including lysine cluster locations predicted in A.

Furthermore, in drop digest using specific proteases (e.g. Trypsin, Chymotrypsin, Subtilisin) can in some cases enhance chances of crystallization (Dong et al. 2001, 2007; Huang et al. 2012; Wernimont and Edwards 2009). Here, the proteases are added to the protein after setting up the drops to remove flexible or unstable structures for improving chances of crystallization. A similar method can be used prior to the screening to optimize constructs. Edman sequencing provides information about the N terminal sequence of protease digestion products that prove to be stable over a longer time period during the digest (Smith 2001) (e.g. after proteolytic removal of flexibel loops). However, since this sequencing method only provides identification of the first few N terminal amino acids of each analyzed digestion fragment *(4 positions with the standard protocol),* this analytic method is only applicable when proteins are *(also)* degraded at their N terminal end. Methods like mass spectrometry can however help with this issue, as peptide coverage can provide information about the respective amino acid composition of stable bands visible in SDS PAGE. Western blot analysis of proteolytic digestions was performed on Uaf30 and the Rrn9/Rrn10 subcomplex, but the results predominantly indicated C terminal digestion, rendering N terminal sequence analysis pointless. Additionally, the digestion patterns, particularly for the Rrn9/Rrn10 subcomplex, exhibited high complexity due to the sample's heterogeneity, as discussed earlier. Based on the limited digest and secondary structure predictions of Uaf30, two C terminal truncations were created. Neither truncation exhibited satisfactory expression levels, and the yields were insufficient for subsequent structural or functional analysis.

5.4 Final thoughts and future perspectives

This work contributes complementary data regarding the function of two UAF subunits. We successfully identified a DNA binding domain within Rrn9, aligning well with the recently published structural data on UAF. We also confirmed a specific binding region for Uaf30 within the upstream element. In context of whole UAF, both of these elements have been demonstrated to play a pivotal role in its functionality. Nevertheless, our findings also spur the emergence of new questions.

One, for example, is the role of Uaf30, in particular its C terminal domain, especially as it could not be resolved in the atomic model and its structure remains undiscovered. This leaves room

for interpretation but also provides an incentive for future Uaf30 related projects. With the basis for structural and functional analyses being established, previously discussed approaches like in drop digest or surface entropy reduction could be suited for structural characterization of this subunit. These could not only be applied to the subunit or domains themselves, but can also be utilized for structural analysis of the whole complex. Especially with regard to the ambiguous orientation of TBP in respect to UAF and Core Factor in the PIC assembly, such approaches could potentially be extended to these transcription factors. Aside from structural analysis, the DNA binding by the C terminal domain of Uaf30 has to be further examined. Here, as discussed, randomization of short stretches inside the rDNA promoter can help to narrow down the exact sequence that is responsible for the sequence specificity of Uaf30. Furthermore, new intriguing questions about the role of Rrn9 arised with our findings in regard to DNA binding specificity, and especially about the effects observed in the in vitro transcription. Follow up assays have to be conducted to clarify the role of the C terminal region, in context of whole UAF in particular. Several approaches for this have been discussed. In addition, new variants have to be created and analyzed to answer questions about the role of the N terminal helix of Rrn9. Single amino acid mutations of the respective interacting residues, as well as point mutations inside the rDNA promoter could help to understand the nature of the interaction of Rrn9 and the promoter. The interfaces that were reported in the publication of Baudin et al. provide a good starting point for these investigations.

The main focus of this project was the characterization of Uaf30 and the Rrn9/10 subcomplex. This was also owed to the fact that expression and purification of Rrn5 could not be achieved. Here, the bacterial UAF expression system could provide solutions. The system can be easily manipulated and expression and purification of mutant complexes are already established. Upcoming projects could use the published model as a basis for new mutants that, for instance, lack the DNA binding domain or contain mutated amino acids of Rrn5 and H3, that are reported to be responsible for DNA contacts. As complex formation has been shown to be possible without each subunit, this could represent a workaround to examine the role of Rrn5 and the proximal H3 in terms of specificity and DNA binding in general.

6 Appendix

6.1 Supplementary data

6.1.1 *Chaetomium thermophilum* homologs

6.1.1.1 Construct overview – Chaetomium thermophilum

As a thermophilic relative of the budding yeast, *Chaetomium thermophilum* has proven to be a good source for homologs that are more likely to crystallize due to their increased thermostability. Two homologs of *S. cerevisiae (S.c.)* UAF subunits could be identified in the *C. thermophilum (C.t.)* genome, using the NCBI and EMBL blast functions specific to the organism (Acland et al. 2013; Bock et al. 2014).

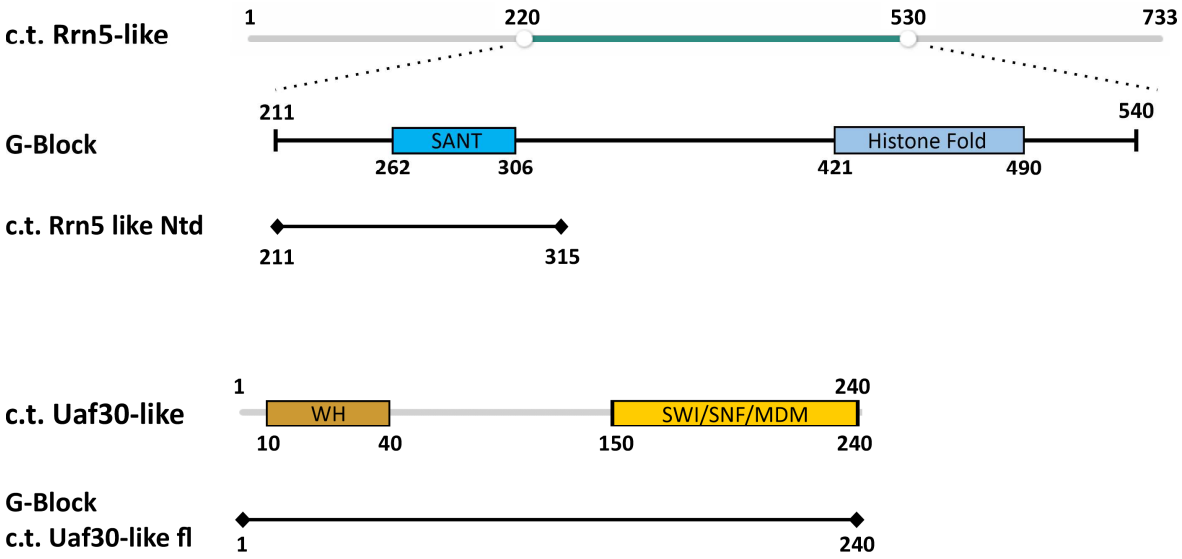
The Rrn5 homolog was identified as a gene that was significantly larger than the gene of its *S.c.* counterpart (gene tag: CTHT_0029410). The two domains that were previously predicted within Rrn5, were located within the middle part of the *C.t.* gene. Based on this finding, a G-Block was ordered, spanning the sequence covering both the SANT and the histone fold domain *(Supplementary Figure 1)*. The sequence of the block was optimized for expression in *E. coli* using the IDT codon usage optimization tool. Several constructs were generated from this block. Only the N terminal region including the SANT domain expressed in *E. coli*, yielding a purification product suitable for crystallization screenings (from here on called: *c.t.* Rrn5-like Ntd).

For Uaf30, a very similar gene was identified (gene tag: CTHT_0000310), carrying both domains that were previously predicted for the *S.c.* UAF subunit. A G-Block spanning the whole gene was ordered and several constructs were generated. Here, only the full length construct expressed well in *E. coli* (from here on called *c.t.* Uaf30-like).

Analyzing the amino acid sequences of both homologs with HHpred after the publication of the structure of UAF, results in two hits that are directly linked to the 2.8 Å UAF structure published by Baudin et al. (2021). This additionally confirms the similarity between the homologs of the two organisms.

Surprisingly, *c.t.* homologs of Rrn9 and Rrn10 could not be identified.

Constructs overview *C.thermophilum*

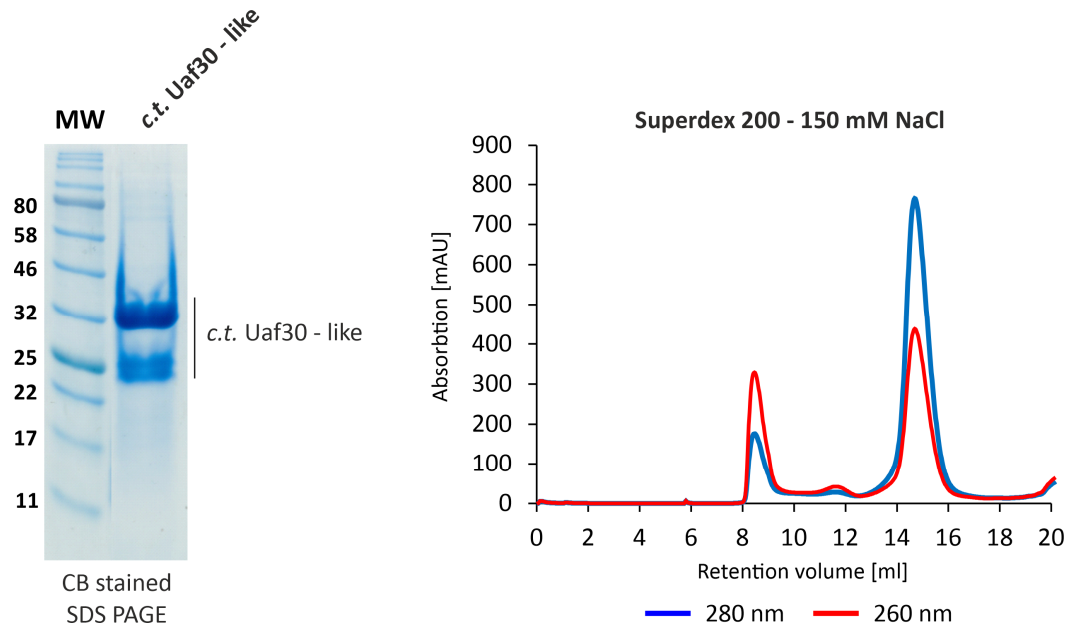


Supplementary Figure 1: Construct overview of *Chaetomium thermophilum* homologs.

An overview of constructs that derived from *C. thermophilum* homologs of yeast UAF subunits that were identified via NCBI or EMBL genome blast. Homologs or Rrn5 (Rrn5-like; genome tag: CTHT_0029410) and Uaf30 (Uaf30-like; genome tag: CTHT_0000310) could be identified. Colored domains were identified using HHpred. The respective domain boundary positions are depicted below the boxes. G-Blocks were created based on stretches with the highest identity to the yeast homologs, yielding several constructs, of which only c.t. Rrn5-like Ntd and c.t. Uaf30-like full length showed expression. WH = winged helix domain

6.1.1.2 Purification of *C.t.* Uaf30-like

C.t. Uaf30-like was purified as described in 3.3.8. The purification yielded ~2 mg / L of highly soluble protein. However, the purification product exhibited strong degradation *(Supplementary Figure 2)*. The degradation pattern was similar to its *S.c.* counterpart, which also partly degraded to a similar molecular weight *(Figure 10),* but to a lesser extend. Although two domains could be identified at the N and C terminal end of the protein, none of the other constructs that were tested for this homolog showed expression. Nevertheless, the protein was subjected to high throughput crystallization screenings at different concentrations *(4.4)*. Unfortunately, no crystallization was observed in any of the tested buffer conditions.

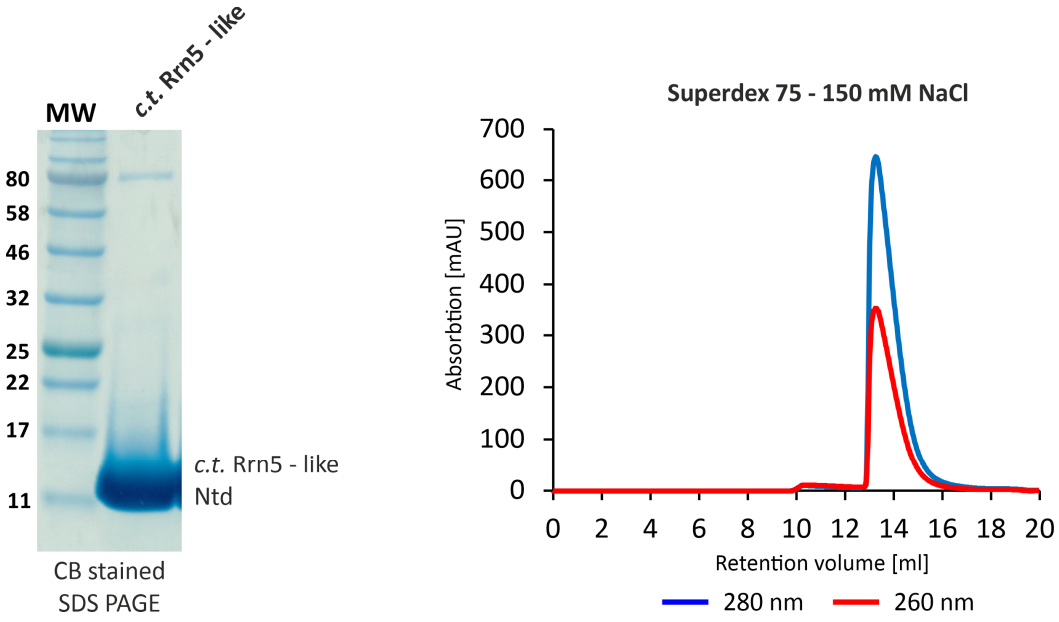


Supplementary Figure 2: Purification of *C.t.* Uaf30-like (size exclusion).

(Left) Coomassie stained SDS PAGE of the *chaetomium thermophilum* homolog of *S.c.* Uaf30: *c.t.* Uaf30-like (SEC eluate). Strong degradations are visible below the main band around 25 kDa. **(right)** Chromatogram of the size exclusion of *c.t.* Uaf30-like, showing a retention volume of ~15 ml (Superdex 200).

6.1.1.3 Purification of *C.t.* Rrn5-like Ntd

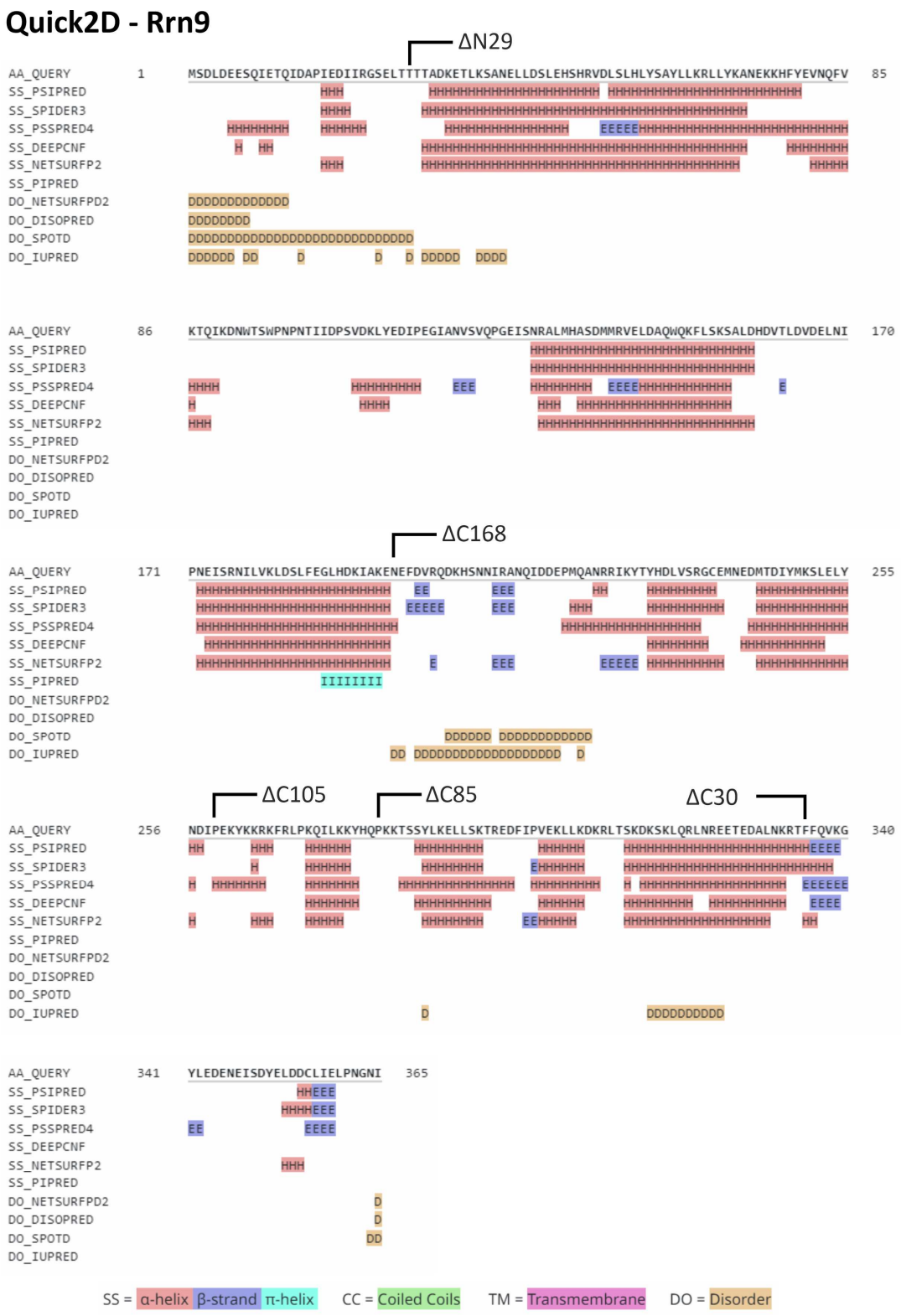
C.t. Rrn5-like Ntd was purified as described in 3.3.9. The purification yielded ~3 mg / L of highly soluble protein *(Supplementary Figure 3)*. A minor contamination was observed at a molecular weight of ~80 kDa, which would correspond to Hsp70, observed for many of the other purifications that were done during this project. Nevertheless, with a purity of > 95 %, the protein was subjected to HT crystallization screenings *(4.4)*. None of the tested concentrations or conditions exhibited any form of crystal formation.



Supplementary Figure 3: Purification of *C.t.* Rrn5-like Ntd (size exclusion).

(Left) Coomassie stained SDS PAGE of the N terminal domain of the *chaetomium thermophilum* homolog of *S.c.* Rrn5: *c.t.* Rrn5-like (SEC eluate). **(right)** Chromatogram of the size exclusion of *c.t.* Rrn5-like, showing a retention volume of ~14 ml (Superdex 75).

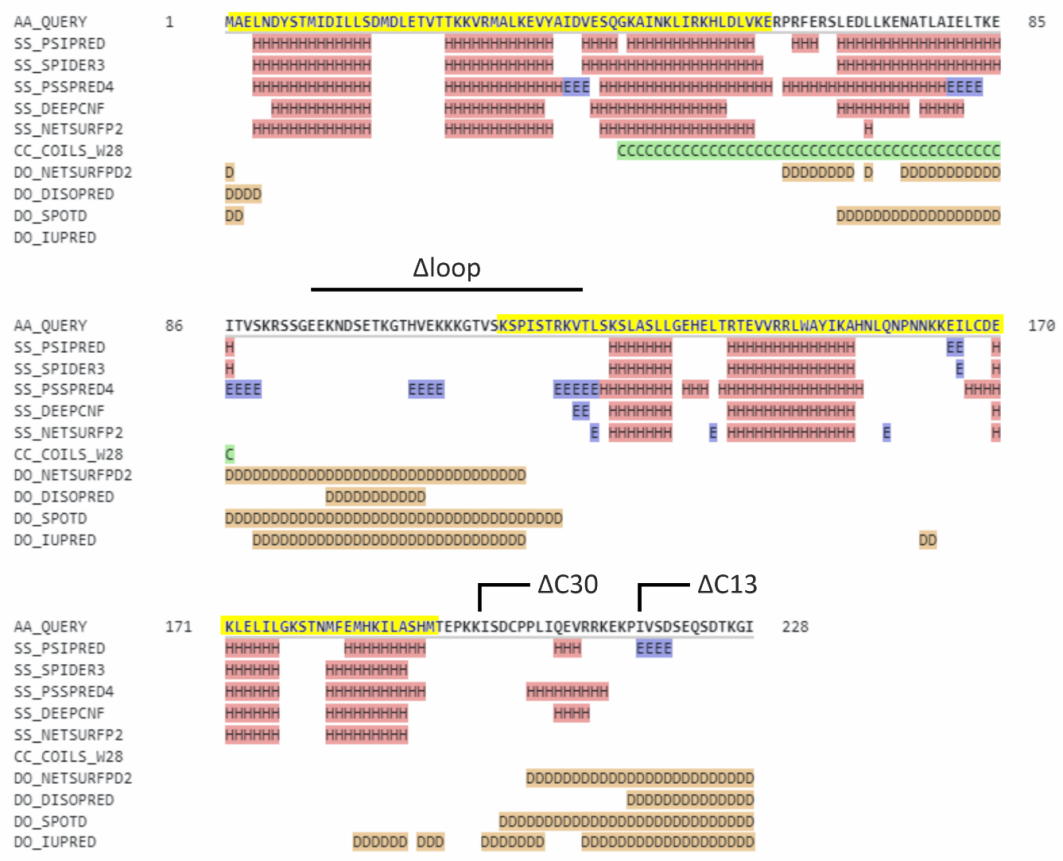
6.1.2 Quick 2D secondary structure predictions



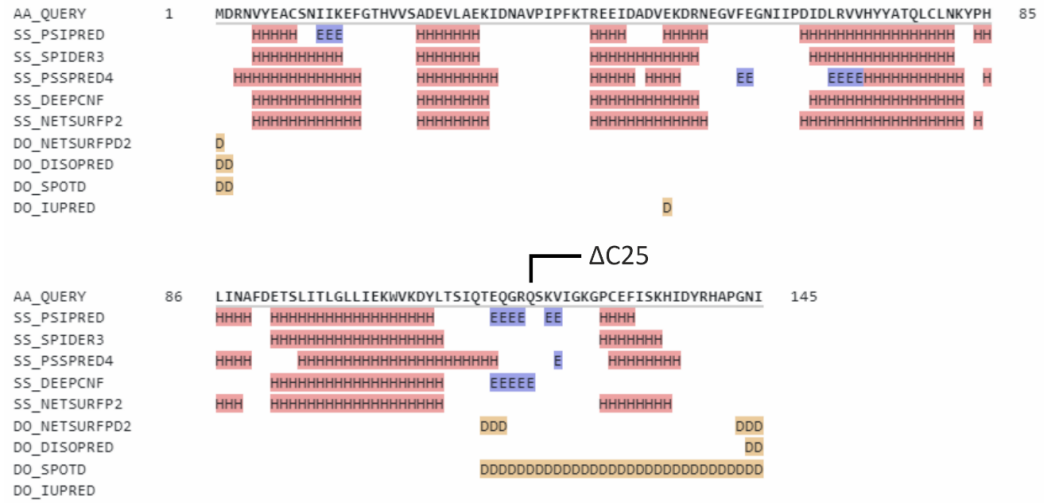
Supplementary Figure 4: Quick 2D - Secondary structure prediction Rrn9.

Overview of secondary structure elements and disordered regions of Rrn9, predicted by the HHpred online server (MPI bioinformatics toolkit). N and C terminal truncations are highlighted. The respective positions show the position of the newly introduced stop codons, which correspond to truncations shown in the construct overview: *Figure 7*

Quick2D - Uaf30



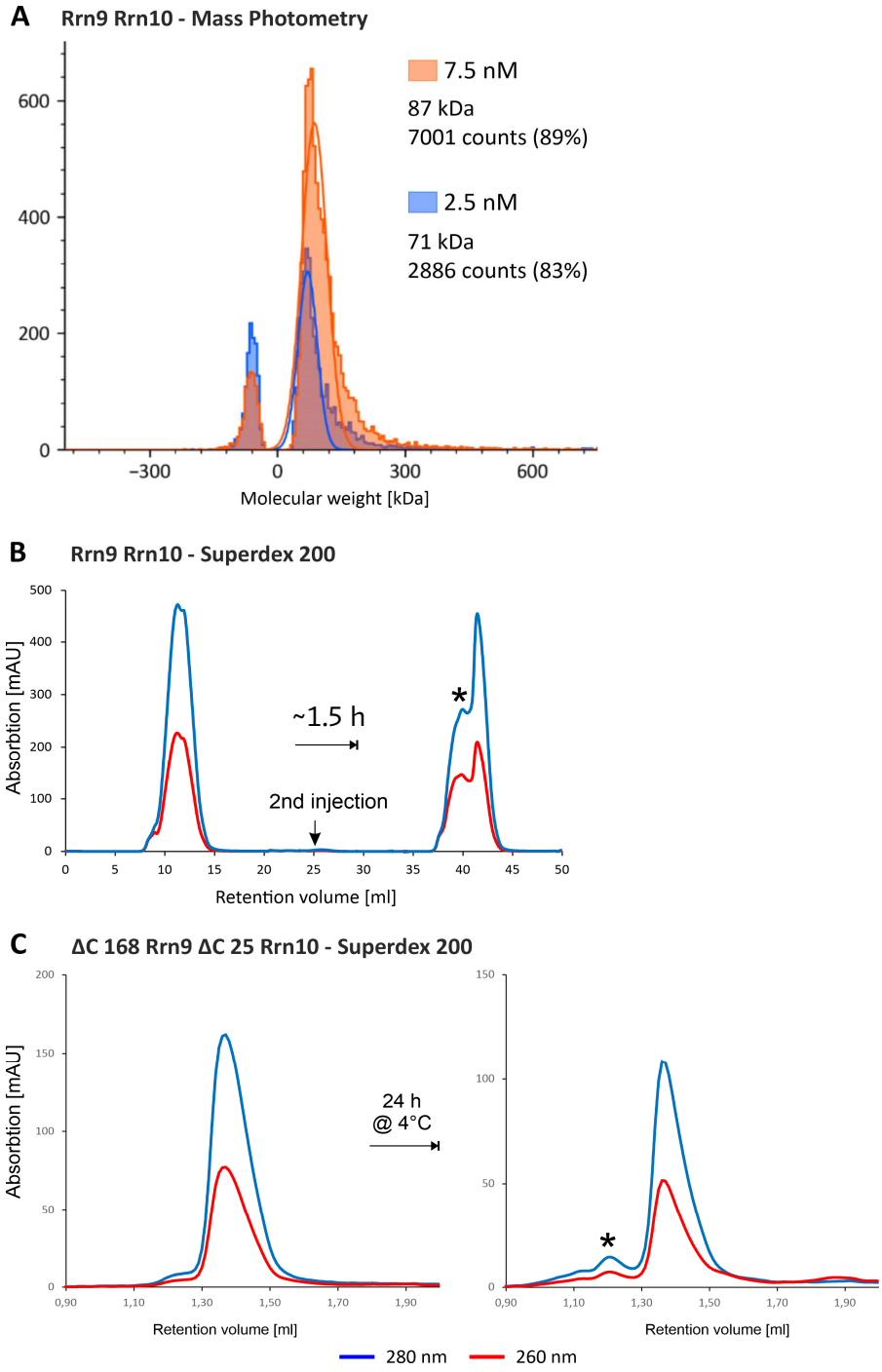
Quick2D - Rrn10



Supplementary Figure 5: Quick 2D - Secondary structure predictions Uaf30 and Rrn10.

Overview of secondary structure elements and disordered regions of Uaf30 (top) and Rrn10 (bottom) predicted by the HHpred online server (MPI bioinformatics toolkit). C terminal truncations are highlighted. The respective positions show the position of the newly introduced stop codons. Uaf30: the terminal domains Uaf30 – Ntd and Uaf30 – Ctd are highlighted in yellow.

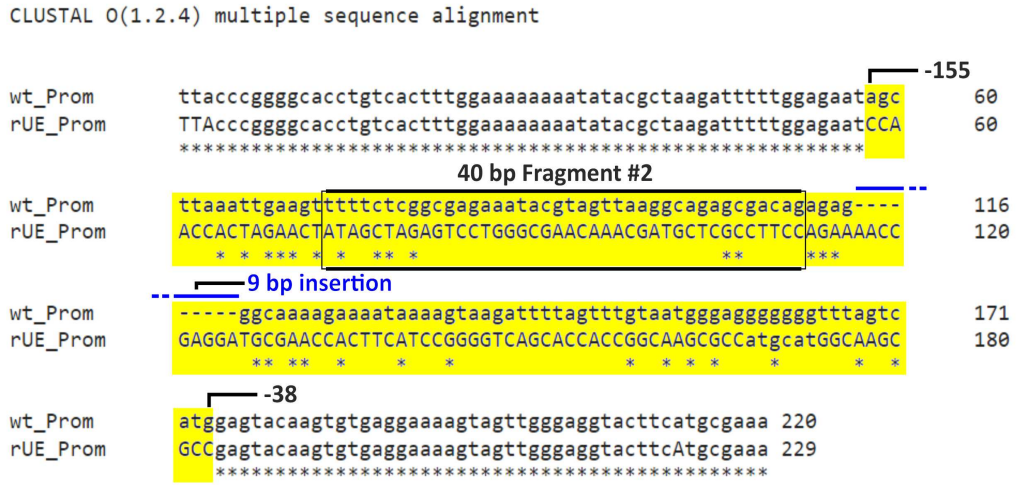
6.1.3 Analysis of Rrn9 / Rrn10 dimerization



Supplementary Figure 6 Rrn9 / Rrn10 Mass Photometry and SEC dimerization analysis.

A) Mass photometry measurements of the Rrn9/10 subcomplex after one freeze/thaw cycle. Two measurements at different concentrations both exhibit a size distribution corresponding to a predominantly monomeric subcomplex. 71 and 87 kDa lie within the error range of the actual molecular weight of ~68 kDa (including the N terminal his tag of Rrn9). **B)** Chromatogram of two consecutive size exclusion (SEC) runs, injected from the same sample. A clear shift towards higher molecular weights is evident in the second run, indicated with an asterisc. **C)** Chromatograms of two SEC runs of a truncated Rrn9/10 subcomplex (ΔC168 Rrn9 ΔC25 Rrn10). The second run was done from the same sample, stored at 4°C for 24 h. While the size distribution still changes, compared to the full length subcomplex, the process is slowed down significantly.

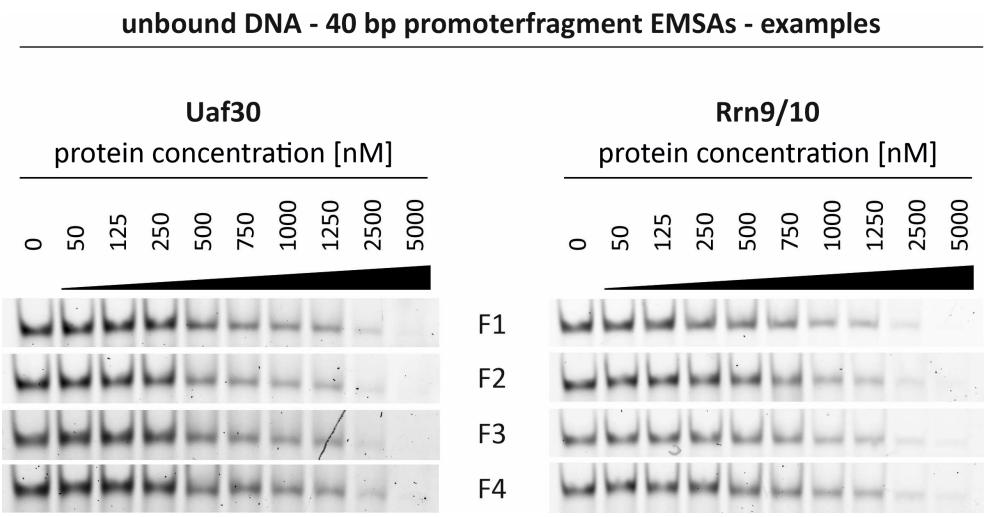
6.1.4 Sequence alignment wildtype rDNA promoter vs randomized UE



Supplementary Figure 7: Sequence alignment of promoter templates for functional analyses

Alignment of the wildtype promoter template and the control template containing a randomized sequence from position -155 to -38 (highlighted in yellow), spanning a length of 220 or 229 bp respectively *(both generated by M. Pilsl)*. A 9 bp insertion was introduced into the control template. Both templates were used in competitive EMSAs and transcription assays, with varying downstream lengths in case of Trx assays. Asteriscs depict identical bases at the given positions. Fragment #2, which exhibited the highest binding affinity towards Uaf30, is highlighted with a black box.

6.1.5 Examplary promoter fragment EMSAs



Supplementary Figure 8: 40 bp promoter fragment EMSAs - Unbound DNA comparison.

Unbound DNA of 40 bp promoterfragment EMSAs of Uaf30 and Rrn9/10 respectively. EMSAs and quantification were done as described in 3.6.1.2. Unbound DNA was quantified with ImageJ, using a rectangle covering only the area visible in the figure for each band. A minimum of 5 runs were done for each concentration. Intensities were normalized to each respective negative control (protein conc. = 0, left lanes). Statisically significant preferential binding was only detected for F2 with Uaf30 *(Figure 17)*.

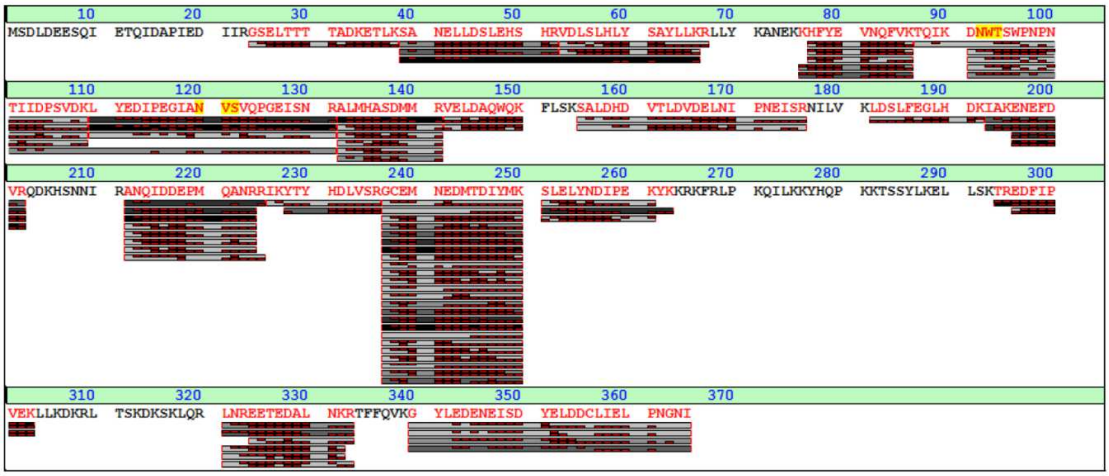
6.1.6 Mass spectrometry - peptide coverages

1 - Rrn9 full length

Results

Protein 1: RNA polymerase I-specific transcription initiation factor RRN9 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) OX=559292 GN=RRN9 PE=1 SV=1

Accession:	sp|P53437|RRN9_YEAST	**Score:**	6128.3
Database:	Hefe_SwissProt_Juni_2018	**Seq. Coverage [%]:**	72.1
MW [kDa] / pI:	42.8 / 5.2	**No. of Peptides:**	77
Modification(s):	Carbamidomethyl, Oxidation, Propionamide, Deamidated		

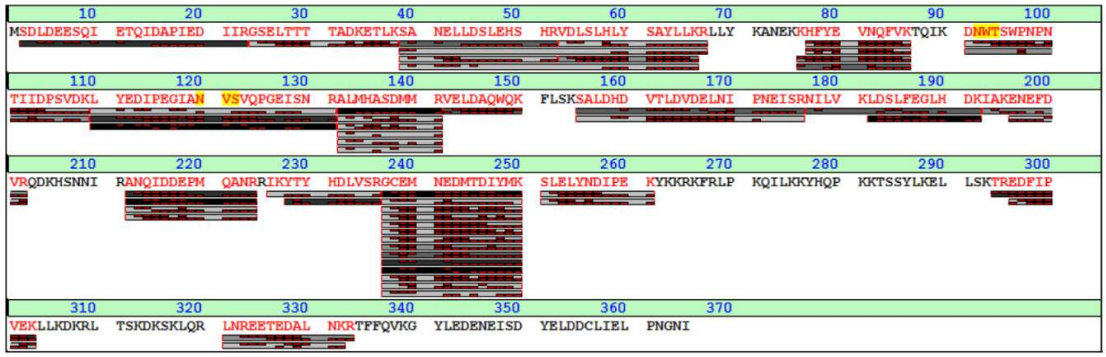


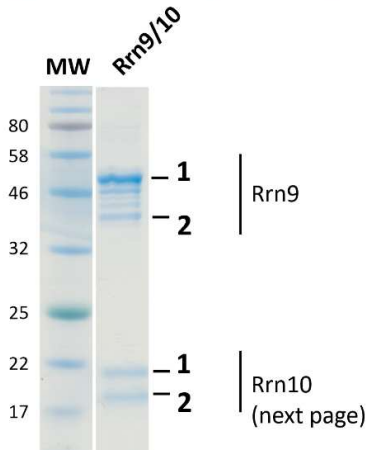
2 - Rrn9 degradation

Results

Protein 1: RNA polymerase I-specific transcription initiation factor RRN9 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) OX=559292 GN=RRN9 PE=1 SV=1

Accession:	sp|P53437|RRN9_YEAST	**Score:**	3880.3
Database:	Hefe_SwissProt_Juni_2018	**Seq. Coverage [%]:**	70.4
MW [kDa] / pI:	42.8 / 5.2	**No. of Peptides:**	55
Modification(s):	Carbamidomethyl, Oxidation, Propionamide, Deamidated		





Supplementary Figure 9: Mass spectrometry peptide coverage of Rrn9.

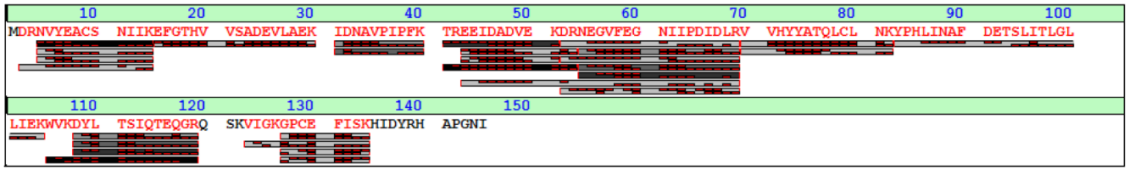
Mass spectrometry peptide coverage of full length Rrn9 (1) and the smallest degradation band (2). The respective samples are indicated on the coomassie stained SDS PAGE (bottom left). Peptide coverage of Rrn10 is shown on the next page. The reduction in peptide coverage of the C terminal region, clearly indicates C terminal degradations, as also evident from the western blot analysis *(Figure 8)*.

1 - Rrn10 full length

Results

Protein 1: RNA polymerase I-specific transcription initiation factor RRN10 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) OX=559292 GN=RRN10 PE=1 SV=1

Accession:	sp|P38204|RRN10_YEAST	**Score:**	2043.7
Database:	Hefe_SwissProt_Juni_2018	**Seq. Coverage [%]:**	89.7
MW [kDa] / pI:	16.5 / 5.1	**No. of Peptides:**	29
Modification(s):	Carbamidomethyl, Propionamide, Deamidated		

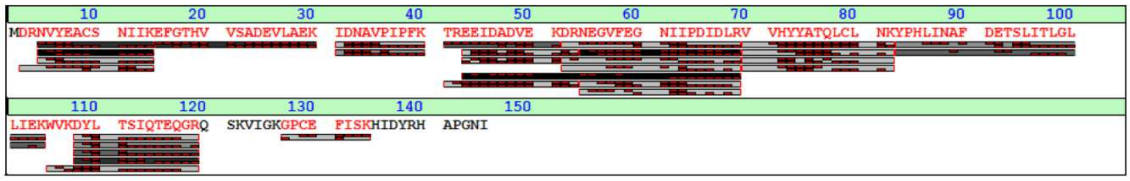


2 - Rrn10 degradation

Results

Protein 1: RNA polymerase I-specific transcription initiation factor RRN10 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) OX=559292 GN=RRN10 PE=1 SV=1

Accession:	sp|P38204|RRN10_YEAST	**Score:**	2032.0
Database:	Hefe_SwissProt_Juni_2018	**Seq. Coverage [%]:**	86.9
MW [kDa] / pI:	16.5 / 5.1	**No. of Peptides:**	29
Modification(s):	Carbamidomethyl, Propionamide, Deamidated		



Supplementary Figure 10: Mass spectrometry peptide coverage of Rrn10.

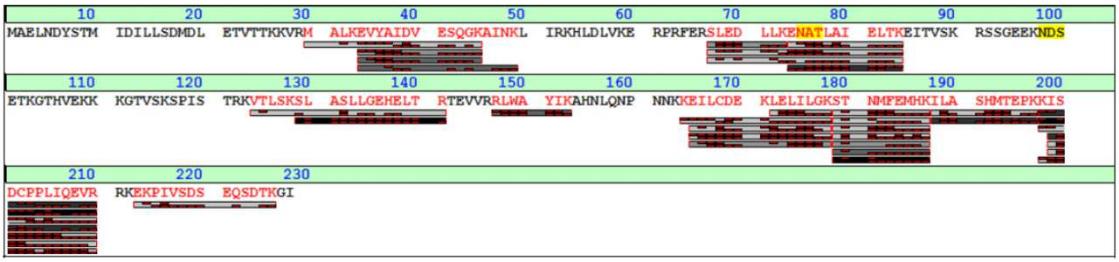
Mass spectrometry peptide coverage of full length Rrn10 (1) and the smallest degradation band (2). The respective samples are indicated on the coomassie stained SDS PAGE (bottom left previous page). A slight decrease in coverage indicates a C terminal degradation of Rrn10.

1 - Uaf30 full length

Results

Protein 1: Upstream activation factor subunit UAF30 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) OX=559292 GN=UAF PE=1 SV=1

Accession:	sp|Q08747|UAF30_YEAST	**Score:**	2157.6
Database:	Swiss_Prot_Hefe_Apr_2020	**Seq. Coverage [%]:**	54.4
MW [kDa] / pI:	26.0 / 8.9	**No. of Peptides:**	34
Modification(s):	Carbamidomethyl, Oxidation, Propionamide, Deamidated		

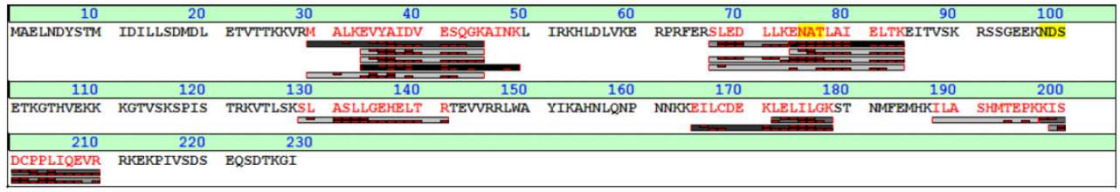


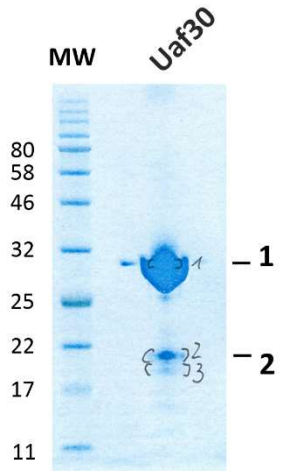
2 - Uaf30 degradation

Results

Protein 1: Upstream activation factor subunit UAF30 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) OX=559292 GN=UAF PE=1 SV=1

Accession:	sp|Q08747|UAF30_YEAST	**Score:**	1012.1
Database:	Swiss_Prot_Hefe_Apr_2020	**Seq. Coverage [%]:**	38.6
MW [kDa] / pI:	26.0 / 8.9	**No. of Peptides:**	16
Modification(s):	Carbamidomethyl, Oxidation, Deamidated		





Supplementary Figure 11: Mass spectrometry peptide coverage of Uaf30.

Mass spectrometry peptide coverage of full length Uaf30 (1) and the its ~20 kDa degradation band (2). The respective samples are indicated on the coomassie stained SDS PAGE (bottom left). A decline of coverage in the C terminal region is evident in #2, indicating C terminal degradation, as also confirmed by western blot *(Figure 10)*.

6.1.7 MatLab script for FA curve fitting

```
% fit anisotropy or intensity data from plate reader

global c;
global number_of_points;
global yA;
global yB;
global yC;
global y;
global constC;
global error;

constC=1; %concentration [nM] of labelled species (constant during experiment)
fileA='UAF30_F1_1'
dataA=load(strcat(fileA,'.txt'));
fileB='UAF30_F1_2'
dataB=load(strcat(fileB,'.txt'));
fileC='UAF30_F1_3'
dataC=load(strcat(fileC,'.txt'));
number_of_datasets=length(dataA(1,:)-1);
c=dataA(:,1); %concentrations
yA=dataA(:,2); %anisotropyA
yB=dataB(:,2); %anisotropyB
yC=dataC(:,2); %anisotropyC


Kd_A=load(sprintf('KdA_UAF30_F1.txt'));
Kd_B=load(sprintf('KdB_UAF30_F1.txt'));
Kd_C=load(sprintf('KdC_UAF30_F1.txt'));

error_Kd=std([Kd_A,Kd_B,Kd_C]);

for i=1:length(yA)
    y(i) = mean([yA(i),yB(i),yC(i)]);
    error_y(i) = std([yA(i),yB(i),yC(i)]);
end

values=[min(y),(max(y)-min(y))/constC,10,constC]; %starting anisotropy, anisotropy
increase, kD

Options = optimset('MaxFunEvals',10000,'MaxIter',100000);
solv=fminsearch('chi_line',values,Options);


%-------------------chi_line-------------------------------------------------
function error = chi_line(x) %returns error for x1,x2
%calculate error in fit
global c;
global y;
global error;

number_of_points=length(c);
error = 0 ;
for ii=1:number_of_points,
        ysim(ii) = (y(ii)-func_binding(x,c(ii)) )^2; %squared difference between
data(y(ii)) and fit
        error=error + ysim(ii); %sums up squared difference
end
```

```
%------------------func_binding-----------------------------------------------

function f=func_binding(x,j) %returns value for A,B at j (=t)
global constC;
A=x(1); % starting value
B=x(2); % amplitude
C=constC; %C= concentration labelled substance (constant)
D=x(3); % kD
f=A+B*((C+j+D)/2-sqrt(((C/2+j/2+D/2).^2)-j*C));


%------------------plot graphs------------------------------------------------

Linear fit
hFig1 = figure(1);
set(hFig1, 'Position',  [100, 100, 800, 600])
hold on; box on;
ii=0:((max(c)-min(c))/100):max(c);
plot(c,y,'o','MarkerFaceColor','r','MarkerEdgeColor','r')
hold on
plot(ii,func_binding(solv,ii),'Color','r');
hold on
errorbar(c,y,error_y,'o','Color','r')
bla=axis;
xlabel('Concentration [nM]')
ylabel('mAnisotropy')
ylim([45,140])
title('UAF30')
fprintf ('kD:   %5.3f nM\n',solv(3));
fprintf ('error:   %5.3f \n',error);
set(findall(gcf,'-property','FontSize'),'FontSize',10,'FontName','Arial');


Semilog fit
hFig2 = figure(2);
set(hFig2, 'Position',  [100, 100, 800, 600])
hold on; box on;
ii=min(c):((max(c)-min(c))/100000):max(c);
plot(c,y,'o','MarkerFaceColor','r','MarkerEdgeColor','r')
set(gca, 'XScale', 'log')
hold on
semilogx(ii,func_binding(solv,ii),'Color','r');
hold on
errorbar(c,y,error_y,'o','Color','r')
bla=axis;
xlabel('Concentration [nM]')
ylabel('mAnisotropy')
xlim([min(c),max(c)])
title('UAF30')
fprintf ('kD:   %5.3f nM\n',solv(3));
fprintf ('error:   %5.3f \n',error_Kd);
set(findall(gcf,'-property','FontSize'),'FontSize',10,'FontName','Arial');
```

6.2 List of figures

6.3 List of tables

6.4 Abbreviations

CE:	Core element
DSB:	Double strand break
EMSA:	Electromobility shift assay
E-pro:	EXP promoters
NTS:	non-transcribed spacer
NGS:	Next generation sequencing
OD:	Optical density
PAGE:	Polyacrylamide gel-electrophoresis
PCR:	Polymerase chain reaction
q-RT PCR:	quantitative real time PCR
Pol I:	RNA polymerase I
PSW:	Polymerase switch (phenotype/strain)
SEC:	Size exclusion chromatography
Trx:	Transcription (assay)
TSS:	Transcription start site
UE:	Upstream element
WB:	Western blot

Abbreviations for protein names are introduced at each respective occasion.

7 References

Acland, Abigail et al. 2013. "Database Resources of the National Center for Biotechnology Information." *Nucleic Acids Research* 41(D1): 8–20.

ALLFREY, V. G., R. FAULKNER, and A. E. MIRSKY. 1964. "Acetylation and Methylation of Histones and Their Possible Role in The." *Proceedings of the National Academy of Sciences of the United States of* 51(1938): 786–94.

Anderson, Susan J. et al. 2011. "The Transcription Elongation Factor Spt5 Influences Transcription by RNA Polymerase I Positively and Negatively." *Journal of Biological Chemistry* 286(21): 18816–24.

Aprikian, Pavel, Beth Moorefield, and Ronald H Reeder. 2000. "TATA Binding Protein Can Stimulate Core-Directed Transcription by Yeast RNA Polymerase I." *Mol Cell Biol* 20(14): 5269–75. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC85976/pdf/mb005269.pdf.

Baudin, Florence et al. 2022. "Mechanism of RNA Polymerase I Selection by Transcription Factor UAF." *Science Advances* 8(16): 1–10.

Bedwell, Gregory J., Francis D. Appling, Susan J. Anderson, and David A. Schneider. 2012. "Efficient Transcription by RNA Polymerase I Using Recombinant Core Factor." *Gene* 492(1): 94–99.

Bell, Stephen P., R. Marc Learned, Hans Michael Jantzen, and Robert Tjian. 1988. "Functional Cooperativity between Transcription Factors UBF1 and SL1 Mediates Human Ribosomal RNA Synthesis." *Science* 241(4870): 1192–98.

Ben-Shem, Adam et al. 2011. "The Structure of the Eukaryotic Ribosome at 3.0 Å Resolution." *Science* 334(6062): 1524–29.

Bier, Mirko, Stephan Fath, and Herbert Tschochner. 2004. "The Composition of the RNA Polymerase I Transcription Machinery Switches from Initiation to Elongation Mode." *FEBS Letters* 564(1–2): 41–46.

Blattner, Claudia et al. 2011. "Molecular Basis of Rrn3-Regulated RNA Polymerase I Initiation and Cell Growth." *Genes and Development* 25(19): 2093–2105.

Bock, Thomas et al. 2014. "An Integrated Approach for Genome Annotation of the Eukaryotic Thermophile Chaetomium Thermophilum." *Nucleic Acids Research* 42(22): 13525–33.

Bordi, L., F. Cioci, and G. Camilloni. 2001. "In Vivo Binding and Hierarchy of Assembly of the Yeast RNA Polymerase I Transcription Factors." *Molecular Biology of the Cell* 12(3): 753–60.

Boukhgalter, Boris et al. 2002. "Characterization of a Fission Yeast Subunit of an RNA Polymerase I Essential

Transcription Initiation Factor, SpRrn7h/TAFI68, That Bridges Yeast and Mammals: Association with SpRrn11h and the Core Ribosomal RNA Gene Promoter." *Gene* 291(1–2): 187–201.

Bradford, Marion M. 1976. "A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding." *Crop Journal* 72(1–2): Pages 248-254.

Braglia, Priscilla et al. 2010. "Role of the RNA/DNA Kinase Grc3 in Transcription Termination by RNA Polymerase I." *EMBO Reports* 11(10): 758–64. http://dx.doi.org/10.1038/embor.2010.130.

Braglia, Priscilla, Junya Kawauchi, and Nick J. Proudfoot. 2011. "Co-Transcriptional RNA Cleavage Provides a Failsafe Termination Mechanism for Yeast RNA Polymerase I." *Nucleic Acids Research* 39(4): 1439–48.

Brewer, Bonita J., Daniel Lockshon, and Walton L. Fangman. 1992. "The Arrest of Replication Forks in the RDNA of Yeast Occurs Independently of Transcription." *Cell* 71(2): 267–76.

Bric, Anka, Catherine A. Radebaugh, and Marvin R. Paule. 2004. "Photocross-Linking of the RNA Polymerase I Preinitiation and Immediate Postinitiation Complexes. Implications for Promoter Recruitment." *Journal of Biological Chemistry* 279(30): 31259–67. http://dx.doi.org/10.1074/jbc.M311828200.

Burkhalter, Martin D., and José M. Sogo. 2004. "RDNA Enhancer Affects Replication Initiation and Mitotic Recombination: Fob1 Mediates Nucleolytic Processing Independently of Replication." *Molecular Cell* 15(3): 409–21.

Butryn, Agata et al. 2015. "Structural Basis for Recognition and Remodeling of the TBP:DNA:NC2 Complex by Mot1." *eLife* 4(AUGUST2015): 1–22.

Cavanaugh, Alice H. et al. 2002. "Rrn3 Phosphorylation Is a Regulatory Checkpoint for Ribosome Biogenesis." *Journal of Biological Chemistry* 277(30): 27423–32. http://dx.doi.org/10.1074/jbc.M201232200.

Cerezo, Emilie et al. 2019. "Maturation of Pre-40S Particles in Yeast and Humans." *Wiley Interdisciplinary Reviews: RNA* 10(1): 1–16.

Cheutin, Thierry et al. 2002. "Three-Dimensional Organization of Active RRNA Genes within the Nucleolus." *Journal of Cell Science* 115(16): 3297–3307.

Choe, Soo Young, Michael C. Schultz, and Ronald H. Reeder. 1992. "In Vitro Definition of the Yeast RNA Polymerase I Promoter." *Nucleic Acids Research* 20(2): 279–85.

Claypool J, French S L, Johzuka K, Eliason K, Vu L, Dodd J A, Beyer A L, Nomura M. 2004. "Tor Pathway Regulates Rrn3p-Dependent Recruitment of Yeast RNA Polymerase I to the Promoter but Does Not Participate in Alteration of the Number of Active Genes." *Molecular Biology of the Cell*.

Cmarko, Dusan, Jana Smigova, Lucia Minichova, and Alexey Popov. 2008. "Nucleolus: The Ribosome Factory."

Histology and Histopathology 23(10): 1291–98.

Conrad-Webb, H, and R A Butow. 1995. "A Polymerase Switch in the Synthesis of RRNA in Saccharomyces Cerevisiae." *Molecular and Cellular Biology* 15(5): 2420–28. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7739526.

Cormack, Brendan P., and Kevin Struhl. 1992. "The TATA-Binding Protein Is Required for Transcription by All Three Nuclear RNA Polymerases in Yeast Cells." *Cell* 69(4): 685–96.

Cramer, P. et al. 2008. "Structure of Eukaryotic RNA Polymerases." *Annual Review of Biophysics* 37: 337–52.

Cramer, Patrick et al. 2000. "Architecture of RNA Polymerase II and Implications for the Transcription Mechanism." *Science* 288(5466): 640–49.

———. 2019. "Organization and Regulation of Gene Transcription." *Nature* 573(7772): 45–54. http://dx.doi.org/10.1038/s41586-019-1517-4.

Dammann, Reinhard, Renzo Lucchini, Theo Koller, and José M. Sogo. 1993. "Chromatin Structures and Transcription of RDNA in Yeast Saccharomyces Cerevisiae." *Nucleic Acids Research* 21(10): 2331–38.

Defossez, Pierre Antoine et al. 1999. "Elimination of Replication Block Protein Fob1 Extends the Life Span of Yeast Mother Cells." *Molecular Cell* 3(4): 447–55.

Deller, Marc C., Leopold Kong, and Bernhard Rupp. 2016. "Protein Stability: A Crystallographer's Perspective." *Acta Crystallographica Section:F Structural Biology Communications* 72: 72–95.

Denissov, Sergey et al. 2007. "Identification of Novel Functional TBP-Binding Sites and General Factor Repertoires." *EMBO Journal* 26(4): 944–54.

Devare, Mayur Nimbadas et al. 2020. "TORC1 Signaling Regulates Cytoplasmic PH through Sir2 in Yeast." *Aging Cell* 19(6): 1–15.

Doerr, Allison. 2006. "Widening the Protein Crystallization Bottleneck." *Nature Methods* 3(12): 961.

Dong, Aiping et al. 2001. "Structure of Human DNMT2, an Enigmatic DNA Methyltransferase Homolog That Displays Denaturant-Resistant Binding to DNA." *Nucleic Acids Research* 29(2): 439–48.

———. 2007. "In Situ Proteolysis for Protein Crystallization and Structure Determination." *Nature Methods* 4(12): 1019–21.

Dörner, Kerstin, Chiara Ruggeri, Ivo Zemp, and Ulrike Kutay. 2023. "Ribosome Biogenesis Factors—from Names to Functions." *The EMBO Journal* 42(7): 1–43.

Drygin, Denis, William G. Rice, and Ingrid Grummt. 2010. "The RNA Polymerase i Transcription Machinery:

An Emerging Target for the Treatment of Cancer." *Annual Review of Pharmacology and Toxicology* 50: 131–56.

Eberharter, Anton, and Peter B. Becker. 2002. "Histone Acetylation: A Switch between Repressive and Permissive Chromatin. Second in Review on Chromatin Dynamics." *EMBO Reports* 3(3): 224–29.

Keener, John, Jonathan A Dodd, Dominique L Alo, and Masayasu Nomura. 1997. "Histones H3 and H4 Are Components of Upstream Activation Factor Required for the High-Level Transcription of Yeast RDNA by RNA Polymerase I." 94(December): 13458–62.

Elion, E A, and J R Warner. 1986. "An RNA Polymerase I Enhancer in Saccharomyces Cerevisiae." *Molecular and Cellular Biology* 6(6): 2089–97.

Elion, Elaine A., and Jonathan R. Warner. 1984. "The Major Promoter Element of RRNA Transcription in Yeast Lies 2 Kb Upstream." *Cell* 39(3 PART 2): 663–73.

Engel, Christoph et al. 2013. "RNA Polymerase i Structure and Transcription Regulation." *Nature* 502(7473): 650–55. http://dx.doi.org/10.1038/nature12712.

———. 2017. "Structural Basis of RNA Polymerase I Transcription Initiation." *Cell* 169(1): 120-131.e22. http://dx.doi.org/10.1016/j.cell.2017.03.003.

Engel, Christoph, Simon Neyer, and Patrick Cramer. 2018. "Distinct Mechanisms of Transcription Initiation by RNA Polymerases I and II." *Annual Review of Biophysics* 47(March): 425–46. https://www.annualreviews.org/doi/10.1146/annurev-biophys-070317-033058.

Fath, Stephan et al. 2001. "Differential Roles of Phosphorylation in the Formation of Transcriptional Active RNA Polymerase I." *Proceedings of the National Academy of Sciences of the United States of America* 98(25): 14334–39.

Fernández-Tornero, Carlos. 2018. "RNA Polymerase I Activation and Hibernation: Unique Mechanisms for Unique Genes." *Transcription* 9(4): 248–54.

Ferreira-Cerca, Sébastien et al. 2005. "Roles of Eukaryotic Ribosomal Proteins in Maturation and Transport of Pre-18S RRNA and Ribosome Function." *Molecular Cell* 20(2): 263–75.

French, Sarah L. et al. 2003. " In Exponentially Growing Saccharomyces Cerevisiae Cells, RRNA Synthesis Is Determined by the Summed RNA Polymerase I Loading Rate Rather than by the Number of Active Genes ." *Molecular and Cellular Biology* 23(5): 1558–68.

Friedrich, J. Karsten et al. 2005. "TBP-TAF Complex SL1 Directs RNA Polymerase I Pre-Initiation Complex Formation and Stabilizes Upstream Binding Factor at the RDNA Promoter." *Journal of Biological Chemistry* 280(33): 29551–58.

Fromont-Racine, Micheline, Bruno Senger, Cosmin Saveanu, and Franco Fasiolo. 2003. "Ribosome Assembly in Eukaryotes." *Gene* 313(1–2): 17–42.

Geiger, Sebastian R, and Sebastian Geiger. 2010. "Structure and Function of RNA Polymerase I Subunits."

Gijsbers, Abril, Takuya Nishigaki, and Nuria Sánchez-Puig. 2016. "Fluorescence Anisotropy as a Tool to Study Protein-Protein Interactions." *Journal of Visualized Experiments* 2016(116): 1–9.

Goetze, H. et al. 2010. "Alternative Chromatin Structures of the 35S RRNA Genes in Saccharomyces Cerevisiae Provide a Molecular Basis for the Selective Recruitment of RNA Polymerases I and II." *Molecular and Cellular Biology* 30(8): 2028–45. http://mcb.asm.org/cgi/doi/10.1128/MCB.01512-09.

Goldschmidt, Lukasz, David R. Cooper, Zygmunt S. Derewenda, and David Eisenberg. 2007. "Toward Rational Protein Crystallization: A Web Server for the Design of Crystallizable Protein Variants." *Protein Science* 16(8): 1569–76.

Gorski, Julia J. et al. 2007. "A Novel TBP-Associated Factor of SL1 Functions in RNA Polymerase I Transcription." *EMBO Journal* 26(6): 1560–68.

Granneman, Sander, and Susan J. Baserga. 2005. "Crosstalk in Gene Expression: Coupling and Co-Regulation of RDNA Transcription, Pre-Ribosome Assembly and Pre-RRNA Processing." *Current Opinion in Cell Biology* 17(3): 281–86.

Grummt, Ingrid, Anne Kuhn, Ingrid Bartsch, and Horst Rosenbauer. 1986. "A Transcription Terminator Located Upstream of the Mouse RDNA Initiation Site Affects RRNA Synthesis." *Cell* 47(6): 901–11.

Gubbey, Tobias Michael. 2017. "Crystal Structure of Core Factor and Its Use to Analyze RNA Polymerase I Initiation."

Haag, Jeremy R., and Craig S. Pikaard. 2007. "RNA Polymerase I: A Multifunctional Molecular Machine." *Cell* 131(7): 1224–25.

Hannig, Katharina et al. 2019. "The C-Terminal Region of Net1 Is an Activator of RNA Polymerase I Transcription with Conserved Features from Yeast to Human." *PLoS Genetics* 15(2).

Heiss, Florian B., Julia L. Daiß, Philipp Becker, and Christoph Engel. 2021. "Conserved Strategies of RNA Polymerase I Hibernation and Activation." *Nature Communications* 12(1): 1–9. http://dx.doi.org/10.1038/s41467-021-21031-8.

Henderson, A. S., D. Warburton, and K. C. Atwood. 1972. "Location of Ribosomal DNA in the Human Chromosome Complement." *Proceedings of the National Academy of Sciences of the United States of America* 69(11): 3394–98.

Henderson, Sheryl, and Barbara Sollner-Webb. 1986. "A Transcriptional Terminator Is a Novel Element of the Promoter of the Mouse Ribosomal RNA Gene." *Cell* 47(6): 891–900.

Herrmann, Helmut, and Herbert Bucksch. 2014. "Crystallization Test." *Dictionary Geotechnical Engineering/Wörterbuch GeoTechnik*: 317–317.

Hinnebusch, Alan G. 2014. "The Scanning Mechanism of Eukaryotic Translation Initiation." *Annual Review of Biochemistry* 83(2): 779–812.

Hontz, R. D. et al. 2008. "Transcription of Multiple Yeast Ribosomal DNA Genes Requires Targeting of UAF to the Promoter by Uaf30." *Molecular and Cellular Biology* 28(21): 6709–19. http://mcb.asm.org/cgi/doi/10.1128/MCB.00703-08.

Hori, Yutaro, Christoph Engel, and Takehiko Kobayashi. 2023. "Regulation of Ribosomal RNA Gene Copy Number, Transcription and Nucleolus Organization in Eukaryotes." *Nature Reviews Molecular Cell Biology*.

Huang, Jinguang et al. 2012. "Seeding with Protease to Optimize Protein Crystallization Conditions in in Situ Proteolysis." *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* 68(5): 606–9.

Iida, Tetsushi, and Takehiko Kobayashi. 2019. "RNA Polymerase I Activators Count and Adjust Ribosomal RNA Gene Copy Number." *Molecular Cell* 73(4): 645-654.e13. https://doi.org/10.1016/j.molcel.2018.11.029.

Jansa, P., and I. Grummt. 1999. "Mechanism of Transcription Termination: PTRF Interacts with the Largest Subunit of RNA Polymerase I and Dissociates Paused Transcription Complexes from Yeast and Mouse." *Molecular and General Genetics* 262(3): 508–14.

Jenner, Lasse et al. 2012. "Crystal Structure of the 80S Yeast Ribosome." *Current Opinion in Structural Biology* 22(6): 759–67. http://dx.doi.org/10.1016/j.sbi.2012.07.013.

Johzuka, Katsuki, and Takashi Horiuchi. 2002. "Replication Fork Block Protein, Fob1, Acts as an RDNA Region Specific Recombinator in S. Cerevisiae." *Genes to Cells* 7(2): 99–113.

Keener, John, Jonathan A. Dodd, Dominique Lalo, and Masayasu Nomura. 1997. "Histones H3 and H4 Are Components of Upstream Activation Factor Required for the High-Level Transcription of Yeast RDNA by RNA Polymerase I." *Proceedings of the National Academy of Sciences of the United States of America* 94(25): 13458–62.

Keener, John, Cathleen A Josaitis, Jonathan A Dodd, and Masayasu Nomura. 1998. "Reconstitution of Yeast RNA Polymerase I Transcription in Vitro." 273(50): 33795–802.

Keys, D. A. et al. 1996. "Multiprotein Transcription Factor UAF Interacts with the Upstream Element of the Yeast RNA Polymerase I Promoter and Forms a Stable Preinitiation Complex." *Genes and Development* 10(7): 887–903.

Keys, Daniel A. et al. 1994. "RRN6 and RRN7 Encode Subunits of a Multiprotein Complex Essential for the Initiation of RDNA Transcription by RNA Polymerase I in Saccharomyces Cerevisiae." *Genes and Development* 8(19): 2349–62.

Kharde, Satyavati et al. 2015. "Symportin 1 Chaperones 5S RNP Assembly during Ribosome Biogenesis by Occupying an Essential RRNA-Binding Site." : 2–9.

Khatter, Heena, Matthias K. Vorländer, and Christoph W. Müller. 2017. "RNA Polymerase I and III: Similar yet Unique." *Current Opinion in Structural Biology* 47: 88–94.

Kim, Minkyu et al. 2004. "The Yeast Rat1 Exonuclease Promotes Transcription Termination by RNA Polymerase II." *Nature* 432(7016): 517–22.

Klinge, Sebastian, and John L. Woolford. 2019. "Ribosome Assembly Coming into Focus." *Nature Reviews Molecular Cell Biology* 20(2): 116–31. http://dx.doi.org/10.1038/s41580-018-0078-y.

Knutson, Bruce A., Marissa L. Smith, Alana E. Belkevich, and Aula M. Fakhouri. 2020. "Molecular Topology of RNA Polymerase I Upstream Activation Factor." *Molecular and Cellular Biology* 40(13): 1–17.

Kobayashi, Takehiko et al. 2004. "SIR2 Regulates Recombination between Different RDNA Repeats, but Not Recombination within Individual RRNA Genes in Yeast Sophila (Hawley and Marcus, 1989). The Phenomenon Called RNA Polymerase Switch in Yeast Also Involves RDNA Repeat Expansion; Yeast Mut." *Cell* 117: 441–53.

———. 2011. "Regulation of Ribosomal RNA Gene Copy Number and Its Role in Modulating Genome Integrity and Evolutionary Adaptability in Yeast." *Cellular and Molecular Life Sciences* 68(8): 1395–1403.

Kobayashi, Takehiko, and Austen R.D. Ganley. 2005. "Molecular Biology: Recombination Regulation by Transcription-Induced Cohesin Dissociation in RDNA Repeats." *Science* 309(5740): 1581–84.

Kobayashi, Takehiko, Denis J. Heck, Masayasu Nomura, and Takashi Horiuchi. 1998. "Expansion and Contraction of Ribosomal DNA Repeats in Saccharomyces Cerevisiae: Requirement of Replication Fork Blocking (Fob1) Protein and the Role of RNA Polymerase I." *Genes and Development* 12(24): 3821–30.

Kressler, Dieter et al. 2012. "Synchronizing Nuclear Import of Ribosomal Proteins with Ribosome Assembly." (November): 666–72.

Kuhn, Claus D. et al. 2007. "Functional Architecture of RNA Polymerase I." *Cell* 131(7): 1260–72.

Kulak, Nils A. et al. 2014. "Minimal, Encapsulated Proteomic-Sample Processing Applied to Copy-Number Estimation in Eukaryotic Cells." *Nature Methods* 11(3): 319–24. http://dx.doi.org/10.1038/nmeth.2834.

Kulkens, Tanja et al. 1991. "The Yeast RNA Polymerase I Promoter: Ribosomal DNA Sequences Involved in Transcription Initiation and Complex Formation in in Vitro." *Nucleic Acids Research* 19(19): 5363–70.

Laferté, Arnaud et al. 2006. "The Transcriptional Activity of RNA Polymerase I Is a Key Determinant for the Level of All Ribosome Components." *Genes and Development* 20(15): 2030–40.

Lalo, Dominique, Joan S. Steffan, Jonathan A. Dodd, and Masayasu Nomura. 1996. "RRN11 Encodes the Third Subunit of the Complex Containing Rrn6p and Rrn7p That Is Essential for the Initiation of RDNA Transcription by Yeast RNA Polymerase I." *Journal of Biological Chemistry* 271(35): 21062–67. http://dx.doi.org/10.1074/jbc.271.35.21062.

Längst, Gernot, Peter B. Becker, and Ingrid Grummt. 1998. "TTF-I Determines the Chromatin Architecture of the Active RDNA Promoter." *EMBO Journal* 17(11): 3135–45.

Leblanc, B., C. Read, and T. Moss. 1993. "Recognition of the Xenopus Ribosomal Core Promoter by the Transcription Factor XUBF Involves Multiple HMG Box Domains and Leads to an XUBF Interdomain Interaction." *EMBO Journal* 12(2): 513–25.

Léger-Silvestre, Isabelle, Stéphanie Trumtel, Jacqueline Noaillac-Depeyre, and Nicole Gas. 1999. "Functional Compartmentalization of the Nucleus in the Budding Yeast Saccharomyces Cerevisiae." *Chromosoma* 108(2): 103–13.

Luger, Karolin et al. 1997. "Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal Structure of the Nucleosome Core Particle at 2.8 A Resolution." *Nature* 389(6648): 251–60.

Mais, Christine et al. 2005. "UBF-Binding Site Arrays Form Pseudo-NORs and Sequester the RNA Polymerase I Transcription Machinery." *Genes and Development* 19(1): 50–64.

Marques, S, and Tiago F. Outeiro. 2013. "Epigenetics: Development and Disease." *Epigenetics: Development and Disease,* 61: 507–25. http://link.springer.com/10.1007/978-94-007-4525-4.

Mason, Stephen W., Eva E. Sander, and Ingrid Grummt. 1997. "Identification of a Transcript Release Activity Acting on Ternary Transcription Complexes Containing Murine RNA Polymerase I." *EMBO Journal* 16(1): 163–72.

Mayer, Christine, Jian Zhao, Xuejun Yuan, and Ingrid Grummt. 2004. "MTOR-Dependent Activation of the Transcription Factor TIF-IA Links RRNA Synthesis to Nutrient Availability." *Genes and Development* 18(4): 423–34.

McClintock, Barbara. 1934. "The Relation of a Particular Chromosomal Element to the Development of the

Nucleoli in Zea Mays." *Zeitschrift für Zellforschung und Mikroskopische Anatomie* 21(2): 294–326.

Merkl, Philipp et al. 2014. " Binding of the Termination Factor Nsi1 to Its Cognate DNA Site Is Sufficient To Terminate RNA Polymerase I Transcription In Vitro and To Induce Termination In Vivo ." *Molecular and Cellular Biology* 34(20): 3817–27.

Merkl, Philipp E. 2013. "Comparative Studies on Elongation and Termination of the Three RNA Polymerases From." *Universität Regensburg*.

Milkereit, Philipp et al. 2001. "Maturation and Intranuclear Transport of Pre-Ribosomes Requires Noc Proteins." *Cell* 105(4): 499–509.

Milkereit, Philipp, and Herbert Tschochner. 1998. "A Specialized Form of RNA Polymerase I, Essential for Initiation and Growth-Dependent Regulation of RRNA Synthesis, Is Disrupted during Transcription." *EMBO Journal* 17(13): 3692–3703.

Miller, Gail et al. 2001. "EMBO J. 2001 Miller." 20(6): 1373–82.

Miller, O. L., and Barbara R. Beatty. 1969. "Visualization of Nucleolar Genes." *Science* 164(3882): 955–57.

Montanaro, Lorenzo, Davide Treré, and Massimo Derenzini. 2008. "Nucleolus, Ribosomes, and Cancer." *American Journal of Pathology* 173(2): 301–10.

Moss, T., F. Langlois, T. Gagnon-Kugler, and V. Stefanovsky. 2007. "A Housekeeper with Power of Attorney: The RRNA Genes in Ribosome Biogenesis." *Cellular and Molecular Life Sciences* 64(1): 29–49.

Moss, T., K. Mitchelson, and R. de Winter. 1985. "The Promotion of Ribosomal Transcription in Eukaryotes." *Oxford surveys on eukaryotic genes* 2(July 2016): 207–50.

Musters, W. et al. 1989. "Linker Scanning of the Yeast RNA Polymerase I Promoter." *Nucleic Acids Research* 17(23): 9661–78.

Németh, Attila. *The Nucleolus Methods and Protocols Methods in Molecular Biology 1455*. http://www.springer.com/series/7651.

Neyer, Simon et al. 2016. "Structure of RNA Polymerase I Transcribing Ribosomal DNA Genes." *Nature* 540(7634): 607–10. http://dx.doi.org/10.1038/nature20561.

Nikitina, T. V., and L. I. Tishchenko. 2005. "RNA Polymerase III Transcription Machinery: Structure and Transcription Regulation." *Molecular Biology* 39(2): 161–72.

Nikolaev, N., O. I. Georgiev, P. V. Venkov, and A. A. Hadjiolov. 1979. "The 37 S Precursor to Ribosomal RNA Is the Primary Transcript of Ribosomal RNA Genes in Saccharomyces Cerevisiae." *Journal of Molecular Biology* 127(3): 297–308.

Nogi, Yasuhisa, Ryoji Yano, and Masayasu Nomura. 1991. "Synthesis of Large RRNAs by RNA Polymerase II in Mutants of Saccharomyces Cerevisiae Defective in RNA Polymerase I." *Proceedings of the National Academy of Sciences of the United States of America* 88(9): 3962–66.

O'Sullivan, Audrey C., Gareth J. Sullivan, and Brian McStay. 2002. "UBF Binding In Vivo Is Not Restricted to Regulatory Sequences within the Vertebrate Ribosomal DNA Repeat." *Molecular and Cellular Biology* 22(2): 657–68.

Oakes, Melanie et al. 1999. "Transcription Factor UAF, Expansion and Contraction of Ribosomal DNA (RDNA) Repeats, and RNA Polymerase Switch in Transcription of Yeast RDNA." *Molecular and Cellular Biology* 19(12): 8559–69. http://mcb.asm.org/lookup/doi/10.1128/MCB.19.12.8559.

Panov, Kostya I., J. Karsten Friedrich, Jackie Russell, and Joost C.B.M. Zomerdijk. 2006. "UBF Activates RNA Polymerase I Transcription by Stimulating Promoter Escape." *EMBO Journal* 25(14): 3310–22.

Papai, Gabor et al. 2020. "Structure of SAGA and Mechanism of TBP Deposition on Gene Promoters." *Nature* 577(7792): 711–16. http://dx.doi.org/10.1038/s41586-020-1944-2.

Peyroche, Gérald et al. 2000. "The Recruitment of RNA Polymerase I on RDNA Is Mediated by the Interaction of the A43 Subunit with Rrn3." *EMBO Journal* 19(20): 5473–82.

Pikaard, C S et al. 1990. "Enhancers for RNA Polymerase I in Mouse Ribosomal DNA." *Molecular and Cellular Biology* 10(9): 4816–25.

Pilsl, Michael. 2021. *Reconstitution of the RNA Polymerase I Initiation Complex from Recombinant Initiation Factors and Regulation of Its Activity*.

Pilsl, Michael, and Christoph Engel. 2020. "Structural Basis of RNA Polymerase I Pre-Initiation Complex Formation and Promoter Melting." *Nature Communications* 11(1): 1–10. http://dx.doi.org/10.1038/s41467-020-15052-y.

Prescott, Elizabeth M. et al. 2004. "Transcriptional Termination by RNA Polymerase I Requires the Small Subunit Rpa12p." *Proceedings of the National Academy of Sciences of the United States of America* 101(16): 6068–73.

Ravarani, Charles N.J. et al. 2020. "Molecular Determinants Underlying Functional Innovations of TBP and Their Impact on Transcription Initiation." *Nature Communications* 11(1): 1–16. http://dx.doi.org/10.1038/s41467-020-16182-z.

Reiter, Alarich et al. 2012. "The Reb1-Homologue Ydr026c/Nsi1 Is Required for Efficient RNA Polymerase i Termination in Yeast." *EMBO Journal* 31(16): 3480–93.

Rowe, L. B., P. M. Janaswami, M. E. Barter, and E. H. Birkenmeier. 1996. "Genetic Mapping of 18S Ribosomal

RNA-Related Loci to Mouse Chromosomes 5, 6, 9, 12, 17, 18, 19, and X." *Mammalian Genome* 7(12): 886–89.

Rudra, Dipayan, and Jonathan R. Warner. 2004. "What Better Measure than Ribosome Synthesis?" *Genes and Development* 18(20): 2431–36.

Russell, Jackie, and Joost C.B.M. Zomerdijk. 2006. "The RNA Polymerase I Transcription Machinery." *Biochemical Society Symposium* 73(73): 203–16.

Sadian, Yashar et al. 2019. "Molecular Insight into RNA Polymerase I Promoter Recognition and Promoter Melting." *Nature Communications* (2019): 1–13. http://dx.doi.org/10.1038/s41467-019-13510-w.

Sanij, Elaine et al. 2008. "UBF Levels Determine the Number of Active Ribosomal RNA Genes in Mammals." *Journal of Cell Biology* 183(7): 1259–74.

Sasano, Yu et al. 2017. "Molecular Breeding of Saccharomyces Cerevisiae with High RNA Content by Harnessing Essential Ribosomal RNA Transcription Regulator." *AMB Express*.

Scandurra, Roberto et al. 1998. "Protein Thermostability in Extremophiles." *Biochimie* 80(11): 933–41.

Scheer, Ulrich, Marc Thiry, and Guy Goessens. 1993. "Structure, Function and Assembly of the Nucleolus." *Trends in Cell Biology* 3(7): 236–41.

Schmeing, T. Martin, and V. Ramakrishnan. 2009. "What Recent Ribosome Structures Have Revealed about the Mechanism of Translation." *Nature* 461(7268): 1234–42. http://dx.doi.org/10.1038/nature08403.

Schnapp, A., and I. Grummt. 1991. "Transcription Complex Formation at the Mouse RDNA Promoter Involves the Stepwise Association of Four Transcription Factors and RNA Polymerase I." *Journal of Biological Chemistry* 266(36): 24588–95. http://dx.doi.org/10.1016/S0021-9258(18)54269-1.

Schneider, David A. et al. 2007. "Transcription Elongation by RNA Polymerase I Is Linked to Efficient RRNA Processing and Ribosome Assembly." *Molecular Cell* 26(2): 217–29.

Schultz, Michael C., Ronald H. Reeder, and Steven Hahn. 1992. "Variants of the TATA-Binding Protein Can Distinguish Subsets of RNA Polymerase I, II, and III Promoters." *Cell* 69(4): 697–702.

Schweizer, E., C. MacKechnie, and H. O. Halvorson. 1969. "The Redundancy of Ribosomal and Transfer RNA Genes in Saccharomyces Cerevisiae." *Journal of Molecular Biology* 40(2): 261–77.

Shaw, Peter, and John Doonan. 2005. "The Nucleolus: Playing by Different Rules?" *Cell Cycle* 4(1): 102–5.

Shou, Wenying et al. 1999. "Exit from Mitosis Is Triggered by Tem1-Dependent Release of the Protein Phosphatase Cdc14 from Nucleolar RENT Complex." *Cell* 97(2): 233–44.

———. 2001. "Net1 Stimulates RNA Polymerase I Transcription and Regulates Nucleolar Structure

Independently of Controlling Mitotic Exit." *Molecular Cell* 8(1): 45–55.

Siddiqi, Imran, John Keener, Loan Vu, and Masayasu Nomura. 2001. "Role of TATA Binding Protein (TBP) in Yeast Ribosomal DNA Transcription by RNA Polymerase I: Defects in the Dual Functions of Transcription Factor UAF Cannot Be Suppressed by TBP." *Molecular and Cellular Biology* 21(7): 2292–97.

Siddiqi, Imran N. et al. 2001. "Transcription of Chromosomal RRNA Genes by Both RNA Polymerase I and II in Yeast Uaf30 Mutants Lacking the 30 KDa Subunit of Transcription Factor UAF." *EMBO Journal* 20(16): 4512–21.

Sledz, Pawel et al. 2010. "New Surface Contacts Formed upon Reductive Lysine Methylation: Improving the Probability of Protein Crystallization." *Protein Science* 19(7): 1395–1404.

Smith, John Bryan. 2001. "Peptide Sequencing by Edman Degradation." *eLS* (1): 1–3.

Smith, Marissa L. et al. 2018. "Reconstitution of RNA Polymerase I Upstream Activating Factor and the Roles of Histones H3 and H4 in Complex Assembly." *Journal of Molecular Biology* 430(5): 641–54. https://doi.org/10.1016/j.jmb.2018.01.003.

Smith, S. David et al. 1990. "Interaction of RNA Polymerase I Transcription Factors with a Promoter in the Nontranscribed Spacer of Rat Ribosomal DNA." *Nucleic Acids Research* 18(7): 1677–1718.

Sonenberg, N, and A G Hinnebusch. 2009. "Regulation of Translation Initiation in Eukaryotes." *Cell* 136(4): 731–45.

Srivastava, A. K., and D. Schlessinger. 1991. "Structure and Organization of Ribosomal DNA." *Biochimie* 73(6): 631–38.

Steffan, Joan S., Daniel A. Keys, Jonathan A. Dodd, and Masayasu Nomura. 1996. "The Role of TBP in RDNA Transcription by RNA Polymerase I in Saccharomyces Cerevisiae: TBP Is Required for Upstream Activation Factor- Dependent Recruitment of Core Factor." *Genes and Development* 10(20): 2551–63.

Steffan, Joan S., Daniel A. Keys, Loan Vu, and Masayasu Nomura. 1998a. " Interaction of TATA-Binding Protein with Upstream Activation Factor Is Required for Activated Transcription of Ribosomal DNA by RNA Polymerase I in Saccharomyces Cerevisiae In Vivo ." *Molecular and Cellular Biology* 18(7): 3752–61.

Steffan, Joan S, Daniel A Keys, Loan Vu, and Masayasu Nomura. 1998b. "Interaction of TATA-Binding Protein with Upstream Activation Factor Is Required for Activated Transcription of Ribosomal DNA by RNA Polymerase I in Saccharomyces Cerevisiae In Vivo." 18(7): 3752–61.

Straight, Aaron F. et al. 1999. "Net1, a Sir2-Associated Nucleolar Protein Required for RDNA Silencing and Nucleolar Integrity 60 Percent of Transcription in Rapidly Growing Yeast Cells * Department of Cell

Biology Cal Role of Silencing at the RDNA Repeats. Gottlieb and Cell 246 Figure 1." *Cell* 97(2): 245–56.

T., Richmond, J. T. Finch, and A. Klug. 1986. "The Structure of the Nucleosome Core Particle." *Biochemical Society Transactions* 14(2): 221–221.

Takeuchi, Yasushi, Takashi Horiuchi, and Takehiko Kobayashi. 2003. "Transcription-Dependent Recombination and the Role of Fork Collision in Yeast RDNA." *Genes and Development* 17(12): 1497–1506.

Thompson, Martin, Rebecca A. Haeusler, Paul D. Good, and David R. Engelke. 2003. "Nucleolar Clustering of Dispersed TRNA Genes." *Science* 302(5649): 1399–1401.

Thomson, Emma, Sébastien Ferreira-Cerca, and Ed Hurt. 2013. "Eukaryotic Ribosome Biogenesis at a Glance." *Journal of Cell Science* 126(21): 4815–21.

Torreira, Eva et al. 2017. "The Dynamic Assembly of Distinct RNA Polymerase i Complexes Modulates RDNA Transcription." *eLife* 6: 1–23.

Trabulo, Sara et al. 2012. "Cell-Penetrating Peptide-Based Systems for Nucleic Acid Delivery: A Biological and Biophysical Approach." *Methods in Enzymology* 509: 277–300.

Trumtel, Stéphanie et al. 2000. "Assembly and Functional Organization of the Nucleolus: Ultrastructural Analysis of Saccharomyces Cerevisiae Mutants." *Molecular Biology of the Cell* 11(6): 2175–89.

Ucuncuoglu, Suleyman et al. 2016. "Direct Characterization of Transcription Elongation by RNA Polymerase I." *PLoS ONE* 11(7): 1–15.

Vannini, Alessandro, and Patrick Cramer. 2012. "Conservation between the RNA Polymerase I, II, and III Transcription Initiation Machineries." *Molecular Cell* 45(4): 439–46. http://dx.doi.org/10.1016/j.molcel.2012.01.023.

Viktorovskaya, Olga V., Francis D. Appling, and David A. Schneider. 2011. "Yeast Transcription Elongation Factor Spt5 Associates with RNA Polymerase I and RNA Polymerase II Directly." *Journal of Biological Chemistry* 286(21): 18825–33.

Vu, L. et al. 1999. "RNA Polymerase Switch in Transcription of Yeast RDNA: Role of Transcription Factor UAF (Upstream Activation Factor) in Silencing RDNA Transcription by RNA Polymerase II." *Proceedings of the National Academy of Sciences* 96(8): 4390–95. http://www.pnas.org/cgi/doi/10.1073/pnas.96.8.4390.

Walter, Thomas S. et al. 2006. "Lysine Methylation as a Routine Rescue Strategy for Protein Crystallization." *Structure* 14(11): 1617–22.

Warner JR. 1999. "The Economics of Ribosome Biosynthesis in Yeast." *Trends in Biochemical Sciences* 24(11): 437–40. https://doi.org/10.1016/S0968-0004(99)01460-7.

Watson, Aleksandra A., and Christopher A. O'Callaghan. 2005. "Crystallization and X-Ray Diffraction Analysis of Human CLEC-2." *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* 61(12): 1094–96.

Wernimont, Amy, and Aled Edwards. 2009. "In Situ Proteolysis to Generate Crystals for Structure Determination: An Update." *PLoS ONE* 4(4).

West, Steven, Natalla Gromak, and Nick J. Proudfoot. 2004. "Human 5′ → 3′ Exonuclease Xm2 Promotes Transcription Termination at Co-Transcriptional Cleavage Sites." *Nature* 432(7016): 522–25.

White, Robert J. 2004. "RNA Polymerase III Transcription and Cancer." *Oncogene* 23(18): 3208–16.

Woolford, John L., and Susan J. Baserga. 2013. "Ribosome Biogenesis in the Yeast Saccharomyces Cerevisiae." *Genetics* 195(3): 643–81.

Zhang, Yinfeng, Martha L. Sikes, Ann L. Beyer, and David A. Schneider. 2009. "The Paf1 Complex Is Required for Efficient Transcription Elongation by RNA Polymerase I." *Proceedings of the National Academy of Sciences of the United States of America* 106(7): 2153–58.

Zhao, Yu, Jung-Hoon Sohn, and Jonathan R. Warner. 2003. "Autoregulation in the Biosynthesis of Ribosomes." *Molecular and Cellular Biology* 23(2): 699–707.

8 Danksagung

Ich möchte hiermit die Gelegenheit nutzen, meine Wertschätzung für einige Personen auszudrücken, die mir auf diesem langen, oft nicht ganz leichten Weg geholfen haben, sowohl auf fachlicher als auch auf persönlicher Ebene.

Allen voran, so wie es sich gehört, dir – Christoph. Ich erinnere mich genau an ein paar deiner Worte bei meinem Vorstellungsgespräch. Ich glaube du wolltest meine Erwartungen dämpfen oder mich schonmal auf das vorbereiten, was kommen könnte:

„aber… Christallographie ist nicht sexy…“

Ich verstehe jetzt, was du meinst… Danke für die Vorwarnung, und danke auch für alles andere. Deine Geduld, deine fachliche Unterstützung, dafür, dass du Bedarf immer erreichbar warst und uns allen viele Freiheiten gelassen hast. Und natürlich für deine Toleranz gegenüber meinen (und Flos) Schlafgewohnheiten! ;) All das weiß ich zu schätzen.

Mein Dank geht außerdem an das fachliche Umfeld ich dem ich arbeiten durfte. Kollegen, die jederzeit ansprechbar waren, gerne mit ihrer fachlichen Expertise ausgeholfen und wegweisende Ideen für mein Projekt geliefert haben: Herbert Tschochner, Joachim Griesenbeck, Michael Pilsl, Philipp Milkereit, Gernot Längst und Gert Bange.
Danke auch an die beiden zusätzlichen Prüfer für meine Verteidigung, Remco Sprangers und Gunter Meister.

And in case some of my fellow PhD students read this: Simone, Kevin, Tiana, Daniela, Michaela, Michi B. & Michi J., Brent, Simon, Alice, Nanni, Fabi (I hope I didn’t forget anyone) … All of you contributed to me really enjoying to come to the lab (most of the time at least). Every day I knew, even if work sometimes sucked and nothing worked, I would still meet great people who would understand, who I could play a round of Kicker with, and who were all in the same boat with me. Thank you for your company, it was much appreciated. I will not forget this time with you all.

Jetzt zum engeren Kreis:

Zuerst möchte ich unserer Labmanagerin **Mona** danken. Wir haben uns direkt gut verstanden, seit du zu uns gekommen bist. Deine (meistens) entspannte und ruhige Art hat es im Labor wirklich erträglich gemacht. Danke, dass du dich darum gekümmert hast, dass

alles rund läuft im Labor. Danke fürs Aushelfen, wenn 2 Hände nicht gereicht haben! Und natürlich danke für das leckere Brot!

Julia – der Kopf des Labors: Danke für deine Hilfe an so vielen Stellen! Danke für deinen fachlichen Input und die Diskussion bei den alltäglichen Problemen im Projekt und im Labor. Ich werde nie wieder vergessen, dass der K_d nicht immer abhängig ist vom Verhältnis zum Liganden! Wer weiß wie lange ich ohne dich gebraucht hätte, das zu bemerken…

Flo, ein besonderer Dank bzw meine besondere Wertschätzung geht an dich:

Was soll ich sagen… What a (long) ride! Bis hierher sind wir 12 Jahre lang (fast) den gleichen Weg gegangen. Danke für die teilweise unglaublich sinnlosen, intensiven Diskussionen über einfach jeden Müll, danke fürs gemeinsame durchstehen all unserer Durststrecken, danke für alles! Wer weiß, ob ich nicht ohne dich irgendwann aufgegeben hätte. Aber, wir haben es geschafft… Ich hoffe du weißt, wie sehr ich deine Gesellschaft als Kollege und als Freund geschätzt habe. Ohne dich wäre es nicht dasselbe gewesen!

Und zu guter Letzt möchte ich von ganzem Herzen meinen Eltern und meiner Familie danken. Ihr seid eine Konstante und die Rückendeckung in meinem Leben, und ohne euch hätte das nicht funktioniert. **Vielen Dank!**