Functional interactions of the Transcription Factor B during transcription initiation in *Pyrococcus furiosus*

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

der Fakultät für Biologie und Vorklinische Medizin

der Universität Regensburg

Vorgelegt von

Stefan Albin Dexl

aus

Neumarkt i.d.OPf.

Regensburg im November 2016

Promotionsgesuch wurde eingereicht am:

18.11.2016

Die Arbeit wurde angeleitet von:

Prof. Dr. Michael Thomm

Unterschrift

Table of contents

I.   Introduction

This work should provide more detailed insights into mechanisms and structural functions of the transcription factor B, and to some extent, the possible role of the replication protein A in the process termed transcription. To investigate the interactions of these two factors the *in vitro* transcription system of the hyperthermophilic organism *Pyrococcus furiosus* was used. The strain was isolated at Porto die Levante, Vulcano, Italy, and described by Fiala, G. and Stetter, K.O (Fiala, Stetter 1986). It belongs to the domain Archaea, which was defined by Woese, Kandler and Wheelis by comparison of the ribosomal RNA (Woese et al. 1990). These studies revealed that basically all living organisms can be referred to one of the three domain of live: Bacteria, Archaea and Eukarya. *Pyrococcus furiosus*, the "rushing fireball", grows optimally under anaerobic conditions at 95°C with a doubling time of 37 minutes, and can use different sugars as carbon source (Fiala, Stetter 1986). In 1996, Hethke et al. established a *Pyrococcus* cell-free transcription system to enable investigation of transcription processes (Hethke et al. 1996). This artificial system allows one to analyze the functions and mechanisms of different transcription factors, as well as the characterization of distinct subunits of the RNA polymerase using an *in vitro* reconstitution approach of this enzyme (Fouqueau et al. 2013). In the following years studies of archaeal and eukaryotic organisms showed similarities in the genomic sequences concerning the transcription apparatus, as well as relationships of transcription regulating proteins between bacterial and archaeal organisms (Kyrpides, Ouzounis 1999). Therefore biochemical analysis of the archaeal transcription system can be useful to reveal evolutionary aspects between the three domains, as well as to make statements for eukaryotic systems concerning function and regulatory mechanisms of the transcription machinery.

The following chapters should give a more detailed insight into the process of transcription, the similarities between transcription machineries in the domains of life, and a detailed functional characterization of the transcription factor B. In addition, a short overview on the replication protein A is given at the end of this introduction, which is also characterized in this thesis.

A.   Transcription - a crucial step in cellular life

Differentiation, cell division, metabolism as well as communication are major events in the life of multicellular organisms. In addition, single-cell organisms also need to response to environmental factors like temperature, nutrients, or toxins for optimal growth. Therefore the regulation of genetic information is a very important step for cells to perform target-driven functions and tasks. Experiments of Oswald Avery, Alfred Hershey and Martha Chase, as well as the discovery of the structure of deoxyribonucleic acid (DNA) by James Watson, Francis Crick and Rosalind Franklin demonstrated that DNA is the central memory of cellular information (Avery et al. 1944; HERSHEY, CHASE 1952; WATSON, Crick 1953). The genetic code is defined as the sequence of the four nucleobases, adenine, cytosine, thymine, and guanine. The so called "Central Dogma" of molecular biology was proclaimed in the late 1950ths and refined in 1970 (Crick 1958, 1970). Herein it was postulated that information derived from DNA is transcribed into RNA, which can serve as a template for protein biosynthesis. The resulting proteins are essential for numerous cellular processes like metabolism, DNA maintaining and repair, signal pathways for cellular response to various stimuli, and many more, which defines the phenotype of an organism. Since the last decades up to today this "Central Dogma" was improved continuously, as new functional classes of RNA molecules and new protein functions, e.g. the reverse transcriptase or post-

transcriptional RNA processing, were discovered (Shapiro 2009; Koonin 2015). Therefore information does not flow only from DNA to RNA to the protein, moreover, a complete and complex network of information flow exists. RNA, in contrast to DNA, contains a reactive OH-species on the second carbon atom at the ribose, and comprises uracil as nucleobase, the demethylated form of thymine. Nowadays a lot of different RNA molecules are known. Beside the well-described classes of transfer RNA (tRNA), messenger RNA (mRNA), and ribosomal RNA (rRNA) a new RNA group of non-coding RNAs was revealed. These RNAs are clustered into small non-coding RNAs (snRNAs), like microRNAs (miRNA), small interfering RNAs (siRNA), Piwi-interacting RNAs (piwiRNA), small nucleolar RNAs (snoRNA) and long non-coding RNAs (lncRNAs) (see reviews (Ghildiyal, Zamore 2009; Bratkovic, Rogelj 2014; Fatica, Bozzoni 2014; Bhartiya, Scaria 2016)).

Despite the large number of RNA molecules with numerous different functions the origin is the same for every type of RNA: they have to be transcribed from DNA. This process is termed transcription and is carried out by large multi-subunit DNA-dependent RNA-polymerase (RNAP) enzymes. Eukaryotic organisms possess up to five RNAPs, and archaea and bacteria have only one enzyme to synthesize RNA, whereas the subunits are homolog to eukaryotic RNAP II (Werner, Grohmann 2011). The eukaryotic RNA-polymerases I - III have specific functions. The RNAP I transcribe only rRNA (Engel et al. 2013), the RNAP II synthesizes mRNA and some small non-coding RNAs (Kornberg 2007), whereas the RNAP III transcribe the 5S rRNA, tRNAs and small non-coding RNAs (Arimbasseri, Maraia 2016). The nuclear RNAP IV and RNAP V are only present in plant species and some algae, they contain 10 or more subunits which are more or less related to subunits of other RNAPs, and are important for small interfering RNA-mediated gene silencing (Landick 2009). To synthesize RNA, the RNAPs have to be recruited to the DNA by interaction with specific transcription factors. These general factors need access to specific sequence motifs, and therefore DNA has to be remodeled first.

1. Genome organization and promoter-DNA accessibility

Transcription is a precisely organized process which enables targeted gene expression, and is regulated by numerous cellular processes. To transcribe a gene specifically transcription factors need access to target DNA sequences. The genetic material, which can comprise millions of base pairs, is structurally organized and condensed by proteins to facilitate compression of the DNA into a single cell.

DNA of eukaryotes is packaged and organized in the nucleus as chromatin, a conglomeration of nucleosomes. A nucleosome consists of a histone protein bound to 145-147bp DNA (Luger et al. 1997). DNA is wrapped around the histones and cannot be the target of transcription factors due to a steric hindrance. Therefore the chromatin structure has to be remodeled in a way that the histones were relocated to expose free DNA. This process is executed in eukaryotic organisms by a large number of proteins which belong to one of four ATP-dependent chromatin remodeling complex families, whereas the histones can also be modified e.g. by acetylation, methylation, phosphorylation or ubiquitination (Witkowski, Foulkes 2015).

Bacteria lack histones or histone-like proteins, and their DNA is packaged as a nucleoid in the cell, whereas the DNA is bound to and organized by nucleoid-associated proteins (NAPs) (Dorman 2014). The most abundant chromatin proteins in bacteria are members of the HU (histone-like protein from *E.coli* strain U93) protein family, and the related protein HTa can also be found in some archaeal species which lack histone-like proteins (Dorman 2009; Zhang et al. 2012b).

Archaeal organisms show different DNA packaging strategies of their nucleoid. The genomic DNA of thermophilic organisms is positively supercoiled as a result of the reverse gyrase enzyme (Brochier-Armanet, Forterre 2007). This enzyme is thought to be exclusive for hyperthermophilic organisms and therefore this DNA conformation is preferred possibly due to an adaption to hot environments (Forterre et al. 1996). In addition, the DNA is further stabilized by DNA-binding proteins. A highly abundant chromatin protein distributed in the archaeal domain is alba (acetylation lowers binding affinity), or proteins of this family, respectively (Laurens et al. 2012). Alba can be modified by acetylation and deacetylation (Wardleworth et al. 2002), whereas *in vitro* experiments revealed that it can condense, bridge and loop DNA, but its *in vivo* dynamics remains unclear (Jelinska et al. 2005; Laurens et al. 2012). In addition to alba, members of the phylum Euryarchaeota possess mainly histone proteins to organize the DNA (Reeve 2003). These proteins are homologous to the eukaryotic H3 and H4 histone subunits and form dimers in solution and tetramers when bound to DNA (Reeve et al. 2004), but lack the typical N- and C-terminal extensions for modifications (Cheung et al. 2000). In contrast, Crenarchaeota lack eukaryotic-like structures, but have own small basic DNA-binding proteins like Cren7, which are highly conserved and exclusive within this phylum, or the related Sul7 proteins (Guo et al. 2008). These chromatin proteins show high similarity to bacterial NAPs (Driessen, Dame 2011). Indeed, genes for eukaryotic-like proteins were also found in some organisms of the Crenarchaeota (Cubonova et al. 2005).

Less is known about the interplay between DNA organizing proteins and transcription factors, which enable recruitment of the RNA polymerase to the promoter site of a gene for RNA synthesis. However, it was shown that if promoter regions are occupied by DNA-binding proteins, the transcription is blocked due to the prevention of factor binding or inhibition of DNA separation (Soares et al. 1998; Xie, Reeve 2004a; Wilkinson et al. 2010). For example, transcription is inhibited in the *M. jannaschii in vitro* system when nucleosome formation at the promoter site occurs (Wilkinson et al. 2010). Similar effects were observed in *M. thermoautotrophicus,* as binding of HMta2 downstream of the transcription start site (TSS) forms a filament that extends to the upstream part of the +1 site, and prevents transcription factor binding (Xie, Reeve 2004a). Interestingly, the same protein does not block the RNA polymerase in the elongation phase, but it lowers the transcription rate (Xie, Reeve 2004a). Global scale analysis revealed that archaeal histones in general are not present at core promoters of archaeal genes and it was shown that the region directly upstream of the TSS is not occupied by histone proteins (Nalabothula et al. 2013). It was pointed out by Peeters et al. that it is more likely in the genome that sequences direct the positioning of nucleosomes to enable binding of transcription factors rather than the transcription factors block the binding of histones in resulting chromatin-free regions (Peeters et al. 2015).

Taken together, it is still enigmatic how transcription is interlinked to genomic organization in archaeal organisms, because the mechanisms of global gene regulation, as well as the goal-driven deposition of chromatin proteins to make DNA accessible for transcription remains to be determined. However, if DNA becomes accessible for transcription factors, numerous proteins, which regulate transcription by repression or activation, interact with the promoter site of the gene.

2. Promoter architecture and regulation of gene expression

Basically two types of promoters are known: core promoters, also known as the single peak or focused promoters, and dispersed or broad peak promoters (Juven-Gershon et al. 2008; Müller et al. 2007) (Figure 1A). Core promoter is defined as a minimal fragment sufficient to direct correct basal levels of transcription initiation by RNAP with a well-defined transcription

start site (TSS) (Butler, Kadonaga 2002; Müller et al. 2007). The broad peak promoters have several start sites distributed over >100 nucleotides and are typically found in CpG islands in vertebrates (Carninci et al. 2006). Both promoter types have specific elements, which serve as interaction platforms for transcription factors. Dispersed promoters lack the TATA-box, downstream promoter element (DPE) and the motif ten element (MTE), which are typical components of core promoters (Juven-Gershon et al. 2008). Furthermore, genes regulated by core promoters are usually issue-specific (Müller et al. 2007), whereas genes regulated by dispersed promoters are mostly ubiquitously expressed (Carninci et al. 2006).

Core promoters often contain the so called TATA-box, also known as Goldberg-Hogness sequence (Sassone-Corsi et al. 1981) (Figure 1B). It is an AT-rich element with the consensus sequence TATAWAAR, whereas the upstream T is most commonly located at -31 or -30 relative to the transcription start site (TTS) +1 (Hausner et al. 1991; Ponjavic et al. 2006; Carninci et al. 2006). This widely used and ancient element is the most conserved promoter motif in archaea and eukaryotes, and is recognized by the general transcription factor TATA binding protein (TBP) (Thomm, Wich 1988; Hausner et al. 1996). Despite the high abundance only 10% of human RNAP II promoters contain a TATA-box (Bajic et al. 2006). A second motif adjacent to the TATA-box is the transcription factor B recognition element (BRE) which is bound by the transcription factor B (TFB) upstream (BRE^u) and/or downstream (BRE^d) the TATA box (Deng, Roberts 2005; Lagrange et al. 1998). The location of the BRE relative to the TATA and the transcription start site defines the transcription direction (Bell et al. 1999). The BRE and the TATA box are strictly required for core promoter dependent transcription, whereas a third element, the Initiator region (Inr) is not (Gehring et al. 2016). This regulatory element encompasses the TSS +1. Sequence alignments of thousands of mammalian transcription start sites showed that the consensus sequence can be restricted to YR, whereas R is the +1 site (Juven-Gershon et al. 2008) and is often an adenine (Butler, Kadonaga 2002). Inr is recognized by the transcription factor IID (TFIID) in eukaryotes and some transcriptional activators in archaea and comprises a high AT content similar to the TATA box (Gehring et al. 2016). This region is often termed the initially melted region (IMR), and can extend up to 12 base pairs upstream the +1 site (Bell et al. 1998), and is an important determinant for the strength of the stimulatory effect of the transcription factor E (TFE) (Blombach et al. 2015). In addition, a proximal promoter element (PPE) exists in archaeal organisms, which is located approximately 10 base pairs upstream of the transcription start site and can increase transcription output through interaction with general transcription factors (GTFs) (Peng et al. 2009). In contrast, in eukaryotic organisms a downstream core promoter element (DPE) can be found 28 to 33 base pairs downstream the TSS, which is important for basal transcription and interacts with the TATA associated factors (TAF) 6 and 9 of the RNAP I system, and TAFII60 and TAFII40 of TFIID of the RNAP II system (Burke, Kadonaga 1996). Promoters containing DPE usually lacks a TATA-box (Müller et al. 2007). Another sequence in eukaryotes was found by computational and biochemical studies and is called the motif ten element (MTE) (Lim et al. 2004). It is located +18 to +27 downstream of the TSS, and, like DPE, functions with the Inr in a cooperative spacer-dependent manner (Lim et al. 2004). Interestingly, optimization of the core promoter elements TATA-box, DPE, MTE, Inr and BRE^d/BRE^u leads to the strongest known *in vitro* promoter (Juven-Gershon et al. 2006). A much more specific promoter region is the so called downstream core element (DCE), which was found in the beta-globin promoter (Lewis et al. 2000) and also characterized in the adeno virus major late promoter (Lee et al. 2005). It consists of three elements S_I (CTTC; from +6 to +11), S_{II} (CTGT; +16 to +21) and S_{III} (AGC; +30 to +34), and occurs with the DPE. A second specific region can be found in

approximately 1% of human core promoters which are TATA-less, and are called X core promoter element 1 (XCPE1). This element is located from -8 to +2 and interacts only with sequence specific activators like NRF1, NF-1 and Sp1 (Tokusumi et al. 2007).
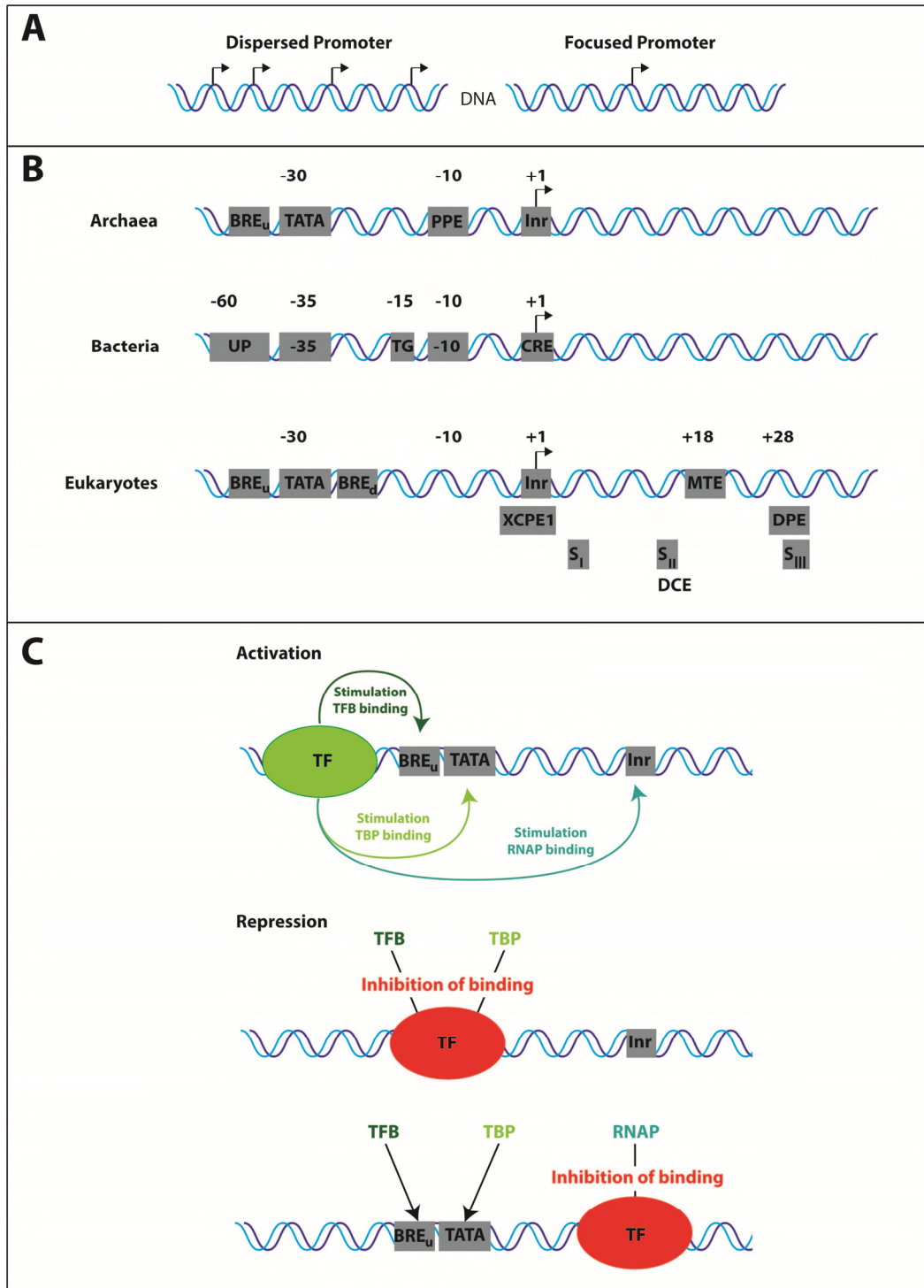
Figure 1: Promoter architecture and regulation of gene expression. **A)** Dispersed and focused (core) promoters differ in the number of their transcription start sites. **B)** General core promoter elements of archaea, bacteria and eukaryotes. **C)** Mechanism of activation and repression of transcription. Transcription factors (TF) bind to sequence motifs upstream the BRE/TATA to activate transcription, whereas binding of TF to elements downstream the BRE/TATA inhibit binding of GTFs and RNAP. (Modified from Peeters, Charlier 2010; Juven-Gershon et al. 2008; Decker, Hinton 2013).

Typical archaeal promoters contain a TATA-box, the BRE and Inr motif. In contrast, Bacteria differ in their promoter architecture in comparison to eukaryotic and archaeal promoters but comprise also sequences important for the interaction with σ-factors and the RNA polymerase. The important sites for interaction with σ-factors are the -35 (TTGACA) and the -10 (TATAAT) region, whereas the AT-rich UP region and the start site containing core recognition element (CRE) both interact with the polymerase (Decker, Hinton 2013). An overview on the common promoter architecture of bacteria, archaea and eukaryotes is shown in figure 1B. The distinct motifs shown here are all *cis*-acting regulatory elements (Butler, Kadonaga 2002), and the presence of distinct motifs and their combinations are one possibility to regulate gene expression (Colgan, Manley 1995). These elements serve as platforms for a variety of transcription factors.

In addition to these combinations gene expression can also be regulated by activators, repressors, enhancers and mediators, which recognize additional specific sequence motifs in proximity to the promoter (Figure 1C). One of the best studied transcriptional regulator in archaea is the Leucine-responsive regulatory protein (Lrp), which possess a typical bacterial helix-turn-helix DNA binding motif, and has a dual role as activator and repressor of transcription (Peeters, Charlier 2010). Members of the Lrp family regulate almost 10% of all genes and are mostly involved in amino acid and central metabolisms in bacteria (Cho et al. 2008). In *Pyrococcus furiosus*, it was shown that the Lrp-like protein LrpA binds closely downstream the TATA box, forming a TBP/TFB/LrpA complex, which in turn blocks the binding of the RNA polymerase due to steric hindrance (Dahlke, Thomm 2002). In contrast, the putative transcription factor 2 (Ptr2) of *Methanococcus jannaschii* activates transcription through binding to an upstream element and stimulates recruitment of TBP (Ouhammouch et al. 2003). A further global regulator of transcription with a dual role is the transcriptional regulator of mal B operon like factor 1 (TrmBL1), which recognizes the *Thermococcales* Glycolytic Motif (TGM) located upstream or downstream of the TATA box to regulate genes involved in sugar metabolism (Gindner et al. 2014). It was shown in ChIP-Seq experiments that TrmBL1 binds to TGMs located downstream of the TATA to repress genes involved in gluconeogenesis, and simultaneously binds to TGMs located upstream of the TATA to switch on genes involved in sugar metabolism under glycolytic growth conditions, whereas TrmBL1 does not bind TGMs under gluconeogenic growth conditions (Reichelt et al. 2016).

The interplay between transcription factors and regulators in combination with distinct promoter elements defines the transcriptional activity and the level of gene expression. The presence of basal factors at the promoter in turn recruits RNAP to initiate RNA-synthesis. Therefore, the gene expression level of a single cell, as a response mechanism to environmental signals, depends on many different factors.

B.   Initiation of transcription: Preinitiation complex formation

The core promoter-dependent transcription process can be divided into three distinct phases. In the first stage general transcription factors specifically interact with sequence motifs of the promoter and bind to DNA until the RNA polymerase is recruited to form a preinitiation complex (PIC). This complex is formed in a stepwise manner as it was shown with native gel electrophoresis experiments (Buratowski et al. 1989) and later with cryo-EM analysis (He et al. 2013). RNAP II preinitiation complexes of eukaryotic organisms consist of in minimum six transcription factors TFIID, TFIIA, TFIIB, TFIIE, TFIIF and TFIIH, whereas archaeal organisms require basically the three eukaryote-related factors, TBP, TFB and TFE (Bell, Jackson 2001; Carlo et al. 2010) (Figure 2). The archaeal transcription machinery therefore constitutes a simplified version of the eukaryotic RNAP II machinery (Grohmann, Werner

2011; Decker, Hinton 2013). In contrast, bacterial complexes contain RNAP and σ (Feng et al. 2016). After complex assembly several structural rearrangements have to take place to convert the initiation complex into an initially transcribing complex. These transitions are shown in chapter I. C (From initiation to elongation and termination). Then RNA synthesis takes place in the elongation phase until transcription is terminated. The proteins which form a preinitiation complex at the core promoter of the three domains are shown in the following chapters.
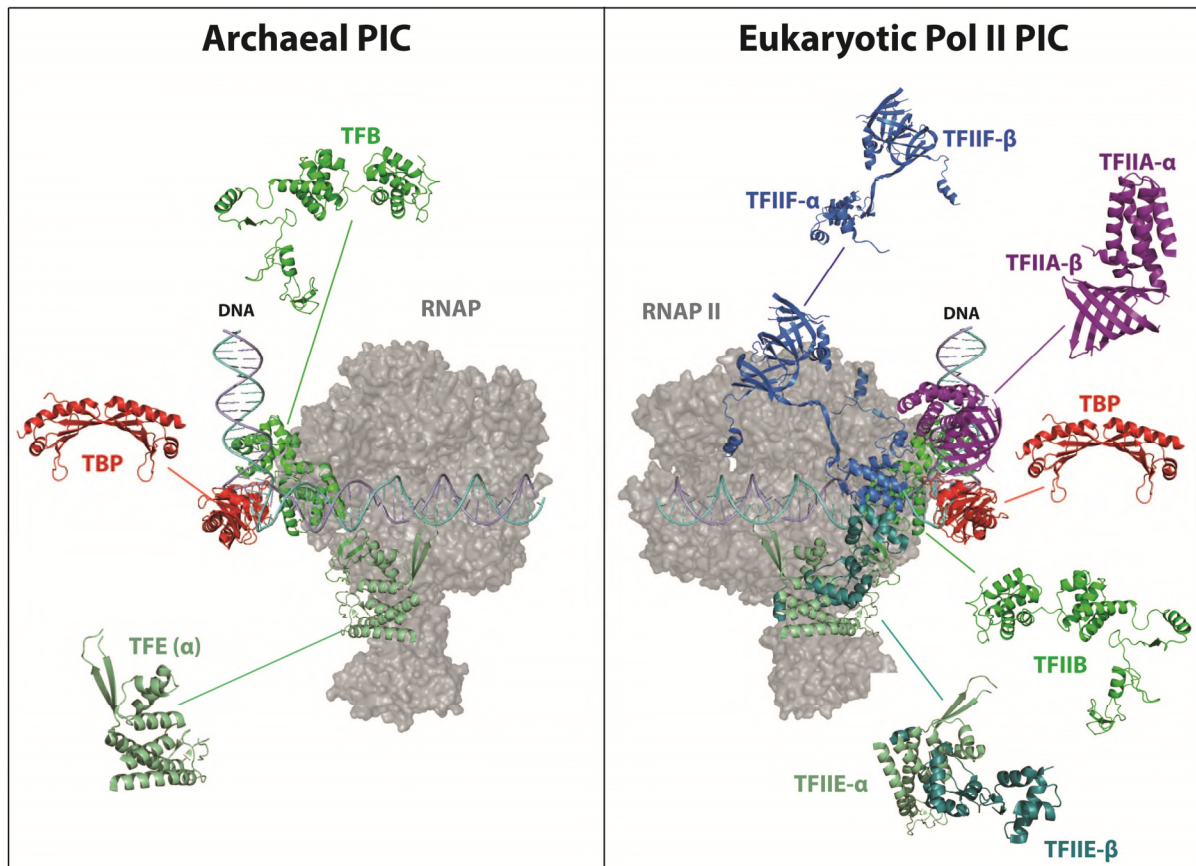
Figure 2: Comparison of archaeal and eukaryotic Pol II preinitiation complexes. Archaeal PIC consists of TBP (red; PDB: 5FZ5), TFB (green; PDB: 3K1F), RNA polymerase (grey; PDB: 4QIW), and TFE (pale green; PDB: 5FZ5) and bent DNA (PDB: 5FZ5), whereas eukaryotic Pol II PIC consist of the related TBP (red; PDB: 5FZ5), TFIIB (green; PDB: 3K1F), TFIIEα/β (pale green and pale blue; PDB: 5FZ5), and the eukaryote-specific TFIIAα/β (purple; PDB: 5FZ5) and TFIIFα/β (blue; PDB: 5FZ5) and bent DNA (PDB: 5FZ5). Complete Pol II PIC structure was modified from PDB: 5FZ5 (Plaschka et al. 2016). For the archaeal PIC *T. kodakarensis* RNAP from structure 4QIW (Jun et al. 2014) was fitted to the complex based on exact overlay of conserved residues in PyMol. TFB/TFIIB was taken from structure 3K1F (Kostrewa et al. 2009) due to absent domain structures in 5ZF5. TFIIH is missing in the 5ZF5 structure because of insufficient resolution of the cryo-EM structure.

1. The TATA binding protein

The first factor which interacts with the TATA element of a core promoter via an induced-fit mechanism is the TATA binding protein (TBP) (Chasman et al. 1993; Kim et al. 1993a; Burley 1996). This protein was formerly referred as the aTFB protein in archaeal organisms, but because of analogous functions to eukaryotic TBP and the similar structure it was re-termed TBP in archaea (Hausner et al. 1996). This saddle-shaped protein comprises a tandem repeat consisting of two conserved domains which are likely the product of ancient gene duplication (Marsh et al. 1994; Adachi et al. 2008). It also possesses an N-terminal extension, which is less conserved and poorly understood in its function (Burley, Roeder

1996). Each of the two domains consists of a five-stranded anti-parallel β-sheet and two α-helices on the opposite site (Kim et al. 1993a; Kim et al. 1993b). Four β-strands of each domain bind to DNA, whereas two α-helices of each domain together with parts of the two β-strands form the convex opposite site and serve as an interface for proteins which are involved in transcription initiation (Akhtar, Veenstra 2011). It was shown that TBP can be exchanged between organisms, e.g. TBP *of P. furiosus* with *Methanococcus* TBP, and *Methanococcus* TBP with human and yeast TBP (Wettach et al. 1995; Hethke et al. 1996).

Bacteria lack the TATA-binding protein and transcription is basically initiated using sigma factors, but it was shown that elements of the conserved TBP are part of the RNase HIII and a DNA glycosylase, likely due to a fusion processes of a TBP core domain and these proteins (Brindefalk et al. 2013). From this point of view, Brindefalk et al. showed that sequences of TBP domains can be found in numerous proteins, indicating that a TBP precursor was present in the last universal common ancestor (LUCA) and evolved either by fusion processes with other proteins or to itself and functions were adapted, or TBP domains originated from DNA-glycosylases and TBP becomes a general transcription factor later (Brindefalk et al. 2013). It is also interesting to note that single TBP-domain sequences were identified e.g. in Halobacteria and in the *Pyrococcus furiosus* genome, which encodes a monopartite TBP of unidentified function in addition to the regular TBP sequence (Brindefalk et al. 2013).

Genomes of higher eukaryotes encode TBP, TBP-related factors (TRF) and TBP-like factors (TLF), which are involved in development and differentiation, in particular gametogenesis and early embryonic development (Akhtar, Veenstra 2011). For RNAP II transcription TBP together with up to 14 TBP associated factors (TAFs) form the eukaryotic TFIID multi-subunit complex (Matangkasombut et al. 2004). The core of TFIID is formed by a subset consisting of TAF4 - TAF6, TAF 8 - TAF10 and TAF12, but no TBP (Leurent et al. 2004). Therefore, different TFIID variations are present in different tissues and cell types to promote targeted gene expression (Demeny et al. 2007), and TBP is further not the universal initiation factor in metazoans like it is in yeast (Akhtar, Veenstra 2011). Recent single molecule analysis on PIC assembly in eukaryotic Pol II transcription showed that TBP alone indeed binds to the promoter, but the specificity of the interaction between TBP and DNA is strongly increased if TBP is part of the TFIID complex (Zhang et al. 2016). In addition, TBP is not only involved in the initiation of RNAP II promoters, it also has its role in RNAP I and RNAP III initiation. TBP together with the selectivity factor SL1 and five RNAP I-specific TAFs are required to initiate RNAP I transcription, whereas TBP and two RNAP III-specific TAFs (BrtI and BdpI) assemble together with TFIIIB to initiate RNAP III transcription (Drygin et al. 2010; Hoffmann et al. 2016).

After TBP bind to the TATA element, the DNA is highly bent in approximately 90° angle due to a transition of DNA into a unique partially unwounded right-handed double helix by a kink (Kim et al. 1993a; Nikolov et al. 1995; Juo et al. 1996). The kink is caused by two phenylalanine residues (Phe284 and Phe301 in human TBP) which contact DNA in the minor groove between the first two base pairs of the TATA-box. The second kink is located at the 7th and 8th base of the TATA-box also by insertion of two phenylalanine residues (Phe193 and Phe210 in human TBP), and DNA is restored back to its usual B-conformation. Binding of TBP and bending of DNA occur simultaneously (Masters et al. 2003) and DNA bending was shown to be a prerequisite for transcriptional activation (Gietl et al. 2014). Recent studies using single molecule analysis showed that DNA in archaeal organisms is bent to one uniform state by TBP alone in the organism *M. jannaschii* of the euryarchaeotic lineage. In contrast, it was shown for *S. acidocaldarius*, a member of the phylum Crenarchaeota, that

the general transcription factor B (TFB) is required to stabilize the bent state (Gietl et al. 2014). For *Saccharomyces cerevisiae* it was shown that DNA bending follows a three-step binding mechanism, as two different complexes were identified with different bending angles, whereas addition of the transcription factor IIB (TFIIB) leads to a fully bent state of the DNA (Gietl et al. 2014).

Beside the stabilization effect of TFIIB, the RNA polymerase II specific auxiliary factor TFIIA also stabilizes the TBP-DNA interaction in eukaryotic transcription initiation (Kang et al. 1995). This factor consists of two conserved domains, a 12-stranded β-barrel which binds to the upstream DNA of the TATA-box and the TBP saddle, and the other domain consists of a four-helix bundle, forming a boot-shaped heterodimer (Tan et al. 1996; Geiger et al. 1996). TFIIA is not able to bind DNA alone, but together with TBP or TFIID, binding to DNA is very efficient (Zhang et al. 2016). TFIIA is not strictly required for transcription initiation, but can stimulate basal and activated transcription (Imbalzano et al. 1994).

2.   The transcription factor B

The next factor associating to DNA and TBP is the general transcription factor B (TFB), or TFIIB for RNAP II transcription, respectively. Recent studies demonstrated that the eukaryotic TFIIB requires TFIID and TFIIA to bind transiently to the promoter, and addition of RNAP II-TFIIF fully stabilizes the association of TFIIB to the preinitiation complex (Zhang et al. 2016). TFB/TFIIB is a single polypeptide consisting of a carboxyl-terminal B-core domain, an amino-terminal Zn-ribbon domain, and a region in between, which were later termed the linker and the reader domain (Ha et al. 1991; Malik et al. 1991; Pinto et al. 1992; Kostrewa et al. 2009) (Figure 3 A). Magnetic resonance spectroscopy analysis of the human TFIIB C-terminal B-core domain (TFIIBc), revealed that this domain consist of two direct repeats which have similar α-helical structures, whereas each repeat contains five alpha-helices A1 to E1 of repeat one, and A2 to E2 of the more hydrophobic repeat 2 (Bagby et al. 1995). First crystal structures of TFIIB/TBP/DNA-complexes indicated that the B-core contacts TBP as well as DNA at the major groove immediately upstream, and at the minor groove downstream the TATA-box (Nikolov et al. 1995). TFB binds to DNA at the BRE via a helix-turn-helix motif formed by helices D and E, whereas the TFIIB-DNA contacts were also verified by DNase I footprinting (Malik et al. 1993), hydroxyl radical footprinting (Lee, Hahn 1995), fluorescence anisotropy measurements and photochemical crosslinking (Lagrange et al. 1998). Mutational analysis of amino acids of the yeast B-core domain further demonstrated these interactions, and revealed that the basic amino acids K190, K201, and K205 play a major role in the interaction with DNA, as these mutants do not form a TBP/TFIIB/DNA complex in yeast in *in vitro* gel shift experiments, and showed impairments in growth *in vivo* (Bangur et al. 1997). Amino acid exchange of the conserved amino acids G153 and R154 (Buratowski, Zhou 1993), as well as amino acid substitutions of G247 and R248 within the second repeat in human TFIIB showed a decreased ability to form TBP/TFIIB/DNA complexes (Bagby et al. 1995). In addition to the mutational analysis, sequence alignments further revealed structural similarities to cyclin A, which is a cell cycle regulating protein (Bagby et al. 1995). From this point of view it was hypothesized that cyclins may have evolved from more fundamental transcription processes in earlier life (Bagby et al. 1995).

Structural analysis of the N-terminus of the archaeal TFB of *Pyrococcus furiosus* showed that this domain forms a Zn-ribbon fold (Zhu et al. 1996). This domain of TFIIB is required for the interaction with the RNA polymerase II associated protein (RAP) 30/74, the small subunit of TFIIF (Ha et al. 1993; Fang, Burton 1996). Moreover, interactions of the Zn-ribbon with RNAP were observed *in vitro* (Tschochner et al. 1992; Bushnell et al. 1996).
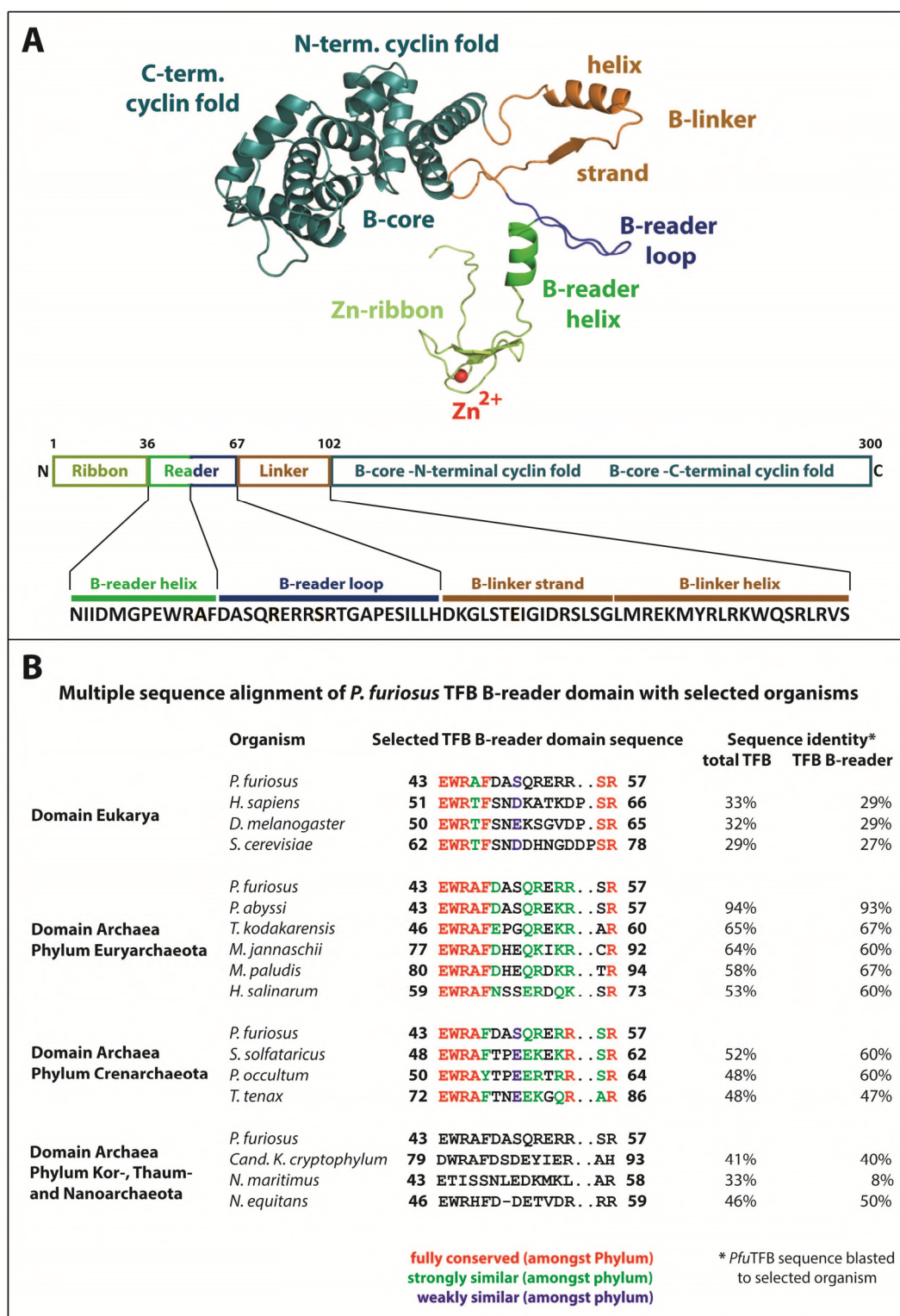
Figure 3: Structure, domain organization and multiple sequence alignments of the transcription factor IIB. **A)** TFB/TFIIB consists of a C-and N-terminal cyclin fold (cyan), a B-linker region (brown), a B-reader domain consisting of the loop (blue) and the helix (green), and a Zn-ribbon (pale green) with a bound Zn^{2+} ion (red) (modified from PDB: 3K1F). The domain organization of *Pfu*TFB is given below from N to the C terminal end, and the same color code is used as in the structure. Amino acids are shown for the respective domain from N36 (reader helix) to S102 (linker helix). **B)** Sequence alignments of the highly conserved TFB/TFIIB B-reader domain. *Pfu*TFB was blasted against respective organisms of different domains and phyla. Identities of the total TFB and the TFB B-reader domain is given as percentages. Amino acids highlighted in red are highly, in green strongly, and in blue weakly conserved within the respective phylum.

RNA polymerase recruitment was shown to be carried out by the Zn-ribbon domain, using amino acid mutagenesis approach in yeast (Pardee et al. 1998). Further site-specific photo crosslinking experiments revealed a specific contact of the TFIIB Zn-ribbon with the surface of the RNAP II dock domain, overlapping the RNA exit point (Chen, Hahn 2003), whereas this location of the ribbon was later confirmed in a crystal structure (Kostrewa et al. 2009). Therefore the Zn-ribbon is essential for RNAP II/TFIIF recruitment. Because of the fact, that TFIIB stabilizes the TBP/DNA complex by direct interactions with TBP and DNA, and the observation, that TFIIB plays a role in RNAP II/TFIIF recruitment, it was proposed that this factor has only a role in bridging between TBP/DNA and RNAP II (Buratowski et al. 1989; Orphanides et al. 1996; Hampsey 1998). Therefore the domains between the N- and C-terminal domains where thought to be just a flexible hinge region. Interestingly, first mutational analysis of the region adjacent to the N-terminal Zn-ribbon revealed that this domain provides key features for the initiation process (Bangur et al. 1997). It was also shown that this region of TFIIB is the most highly conserved region amongst known TFB proteins (Na, Hampsey 1993) (Figure 3 B), and amino acids 52-140 of yeast TFIIB can be functionally replaced by the corresponding region of human TFIIB (Shaw et al. 1996).

The TFB B-reader helix domain is important for transcription start site selection. Mutational analysis of this domain in yeast, especially amino acid R64, results in shifts of the transcription start site *in vitro* and *in vivo*, a cold-sensitive phenotype and diminished growth rates (Pardee et al. 1998; Bangur et al. 1997; Pinto et al. 1992). Amino acid E62 of yeast TFIIB showed the same effects, but interestingly, the corresponding amino acid E51 of human TFIIB is not affected by substitutions (Cho, Buratowski 1999). Therefore it was assumed that the transcription start site selection also depends on the distance between Inr and TATA, because in yeast promoters the spacing between Inr and TATA differ in comparison to human promoters (Cho, Buratowski 1999). Additional analysis of TFIIB B-reader mutations and different Inr sequences made clear that TSS selection is B-reader helix and RNA polymerase dependent (Li et al. 1994; Faitar et al. 2001). Moreover, it was postulated that the TSS selection is carried out by scanning of the RNA polymerase to search for the correct nucleotide to start RNA synthesis (Giardina, Lis 1993). First models of crystal structures containing TFIIB and RNAP II suggest that the B-reader domain might contact one strand of the DNA, indicating a supporting role for TSS selection (Bushnell et al. 2004). In addition, mutational analysis of the upstream region immediately next to the Inr site, especially position -8 eight nucleotides upstream the TSS, in combination with TFIIB B-reader helix mutations also showed altered patterns in the TSS selection (Kuehner, Brow 2006). In a later published model of a yeast TFIIB/RNAP II crystal structure a contact of TFIIB B-reader helix and DNA eight nucleotides upstream the TSS was proposed, which confirmed previous results and strengthen the DNA scanning hypothesis (Kostrewa et al. 2009).

One of the first biochemical analyses of the B-linker domain of the archaeal *Pyrococcus furiosus* TFB revealed that mutations or deletion of this domain indeed form a preinitiation complex, but transcriptional activity is completely lost, and promoter DNA is not melted anymore, indicating that this domain of TFB plays a key role in promoter opening (Kostrewa et al. 2009).

The last domain to mention is the TFIIB B-reader loop domain. Because of the close proximity of the B-reader to the active site of the RNAP in crystal structures containing yeast RNAP II, TFIIB and DNA it was hypothesized that the B-reader loop stabilizes the transcription bubble (Bushnell et al. 2004), and the separation of RNA from DNA was proposed to be carried out by the loop domain by charge-repulsion (Sainsbury et al. 2013).

Beside the above-mentioned functions of the transcription factor II B, it was also shown that this factor can be the target for several transcription regulation factors to activate or repress transcription. It was shown in affinity chromatography experiments that members of the Jun activator protein family can directly interact with the B-core domain (Franklin et al. 1995). The receptor for the thyroxine hormone in chicken (cTR3) was shown to bind efficiently TFIIB in *in vitro* binding studies (Hadzic et al. 1995), and in a yeast two-hybrid protein interaction assay, a specific protein-protein interaction between TFIIB and the vitamin D receptor was shown (MacDonald et al. 1995). In addition, Krüppel, a segmentation protein in Drosophila, also interacts with TFIIB when bound to DNA, and activates transcription (Sauer et al. 1995). Another example of specific gene regulation is the cAMP-induced transcription of cAMP-controlled genes. Here, the c̲AMP r̲esponsive e̲lement b̲inding protein (CREB) can independently and specifically interact with TFIIB in co-immunoprecipitation assays (Xing et al. 1995). These few examples show that, beside the crucial function in basal transcription initiation, TFIIB can also be the target for transcriptional regulators.

Orthologues of TFIIB exist in the transcription system of RNA polymerase I and RNA polymerase III. For RNAP I a TFIIB-like protein was not observed, but with structural predictions based on computational analysis of specific domains of TFIIB, a factor was identified, which comprise the cyclin-folds of the B-core domain, the Zn-ribbon domain and a hinge region similar to the B-reader and B-liker domain (Naidu et al. 2011). This protein is a TBP-associated factor 1B (TAF1B) in human, and is a subunit of the transcription factor SL1. It was shown that TAF1B interacts with the RNAP I recruitment factor hRRN3, which converts the inactive Pol I to an initiation-competent enzyme (Engel et al. 2016), and therefore plays a role in the recruitment of initiation-competent RNA polymerase I to the rDNA promoter (Miller et al. 2001). TAF1B lacks the highly conserved B-reader and B-linker region, and the Zn-ribbon domain plays a role in post-recruitment of the RNAP I in humans (Naidu et al. 2011). It is also interesting to note that the yeast counterpart of TAF1B, Rrn7, has little homology to TAF1B, suggesting a co-evolution of the two factors with species-specific elements of RNAP I (Naidu et al. 2011).

Transcriptional activity and recruitment of RNAP III requires the transcription factor IIIB. This factor is placed on TATA-less promoters of the Pol III system by TFIIIC, or TFIIIB can autonomously interact with the few TATA-boxes present in the RNAP III system (Dieci et al. 2000). Once TFIIIB is bound to the promoter it repetitively recruits the 17 subunit RNAP III (Kassavetis et al. 1990). TFIIIB consists of three subunits, a TATA binding protein, a TFIIB̲-r̲elated f̲actor 1̲ (Brf1) and the RNAP III specific B̲ d̲ouble p̲rime 1̲ (Bdp1), whereas human RNAP III contains two homologous Brf proteins, hsBrf1 and hsBrf2 (Willis 2002). The N-terminus of Brf1 comprises the Zn-ribbon structure of the corresponding TFIIB, but this domain is not essential for the recruitment of RNAP III (Kassavetis, Geiduschek 2006). It was shown that Brf1 and Bdp1 mutations failed to open the promoter and therefore the N-terminal Brf1 domain likely stabilizes the transcribed strand after DNA melting and is essential for TFIIIB activity (Kassavetis et al. 2001).

Bacteria basically lack TFB and their transcription is initiated by σ-factors. Different types of σ-factors evolved in bacterial organisms to regulate targeted gene expression. These proteins can be classified into two major groups. The housekeeping factors or σ^{70}, are necessary for transcription of genes important for cell growth, and can be further classified into group 1-4, whereas the members of these groups differ in absence or presence of four distinct σ-domains (σR1.1, σR1.2-2.4, σR3.0-3.2 and σR4.1-4.2 (Lonetto et al. 1992). The second major group comprises the unrelated σ^{54} factors, which are involved in gene regulation as a response to environmental signals, and require ATP (Paget 2015). For

example one prominent member of the σ^{54} family is the well described σ^{32} factor, the product of the *htpR* gene in *E.coli*, which is important for heat shock response and regulation (Grossman et al. 1984). The domains of σ^{70} factors are thought to interact with each other to maintain a relatively compact organization, making them unable to recognize promoter DNA, but this conformation is changed if the factor associates with the RNA polymerase (Callaci et al. 1998). The σ factors bind to the bacterial core RNAP ($\alpha_2\beta\beta'\omega$) to form a holoenzyme ($\alpha_2\beta\beta'\omega\sigma$) which is recruited to a bacterial promoter to initiate transcription via interaction of σ region 2 with the -10 promoter element, and σ region 4 with the -35 promoter element (Feklistov et al. 2014). TFB/TFIIB and σ-factors differ in their structure and the general domain composition. However, sigma factors were shown to have similar tasks in transcription initiation, like direction of the RNA polymerase to the transcription start site and the support of DNA melting and separation of the DNA strands (Feklistov et al. 2014). In structures of eukaryotic and bacterial preinitiation complexes some regions of the σ-factor likely contact the same regions of DNA and the RNAP as the corresponding domains in TFB/TFIIB. The B-finger, which consists of the B-reader helix and the B-reader loop domain of TFB/TFIIB and the σ-finger of the conserved linker R3.2 between the domains $\sigma3$ and $\sigma4$ are closely located to the transcribed strand of the active site of the RNA polymerase, indicating direct and similar roles in initiation (Liu et al. 2010; Zhang et al. 2012a; Sainsbury et al. 2013). Recent structural comparisons and sequence alignments between archaeal TFB, eukaryotic TFIIB and bacterial σ-factors revealed that these factors are homologues (Burton, Burton 2014). Burton and Burton pointed out that a primordial initiation factor was present in the LUCA comprising more elements of σ, and this factor radiates as the two lineages Archaea and Bacteria arise, whereas in Bacteria the σ co-evolves with the RNAP and the promoter sequence motif -35. Archaeal TFB has lost two HTH motifs and gained a Zn-ribbon and a conserved B-reader element, which then evolved in both, the eukaryotic and the archaeal lineage (Burton, Burton 2014).

It was already observed earlier, that the eukaryotic TFIIB and the archaeal TFB are highly homologous in structure, domain composition and function (Thomm 1996). Interestingly, some archaeal species comprise multiple TFB proteins with different functions, e.g. in *Haloferax volcanii*, several TFB proteins were identified (Thompson et al. 1999). The genome of *Halobacterium salinarum* NRC1 strain encodes seven TFB proteins, whereas tfbF was shown to be essential for growth under standard conditions (Facciotti et al. 2007), and tfbA, tfbC, and tfbG are not (Coker, DasSarma 2007). These proteins function together, as an interaction between the proteins among themselves were demonstrated in protein-protein interaction assays (Facciotti et al. 2007). TfbF further plays a role in temperature-dependent gene regulation (Bleiholder et al. 2012). *Sulfolobus solfataricus* and related members of the phylum Crenarchaeota contain three paralogs of TFB, whereas tfb3 lack the B-finger and DNA binding domains, but is significantly upregulated under UV-light exposure (Paytubi, White 2009). In the genome of the used model organism *Pyrococcus furiosus*, two TFB proteins are encoded, whereas TFB1 is homologous to eukaryotic TFIIB, and TFB2 was proposed to be a paralog of TFB1 (Micorescu et al. 2008). TFB2 lacks the typical conserved B-finger motif in comparison to TFB1, and it was suggested that TFB2 is expressed under heat-shock conditions (Shockley et al. 2003).

Taken together, the transcription factor B and its related proteins TFIIB, σ, Brf1, Rrn7 and the multiple TFB copies of some archaea are present under different conditions at different RNAP systems to initiate transcription. Beside the various functions in regulation of gene expression of the TFB variants, the role of the general transcription factor is to recruit the RNA polymerase to the promoter and support DNA melting and TSS selection.
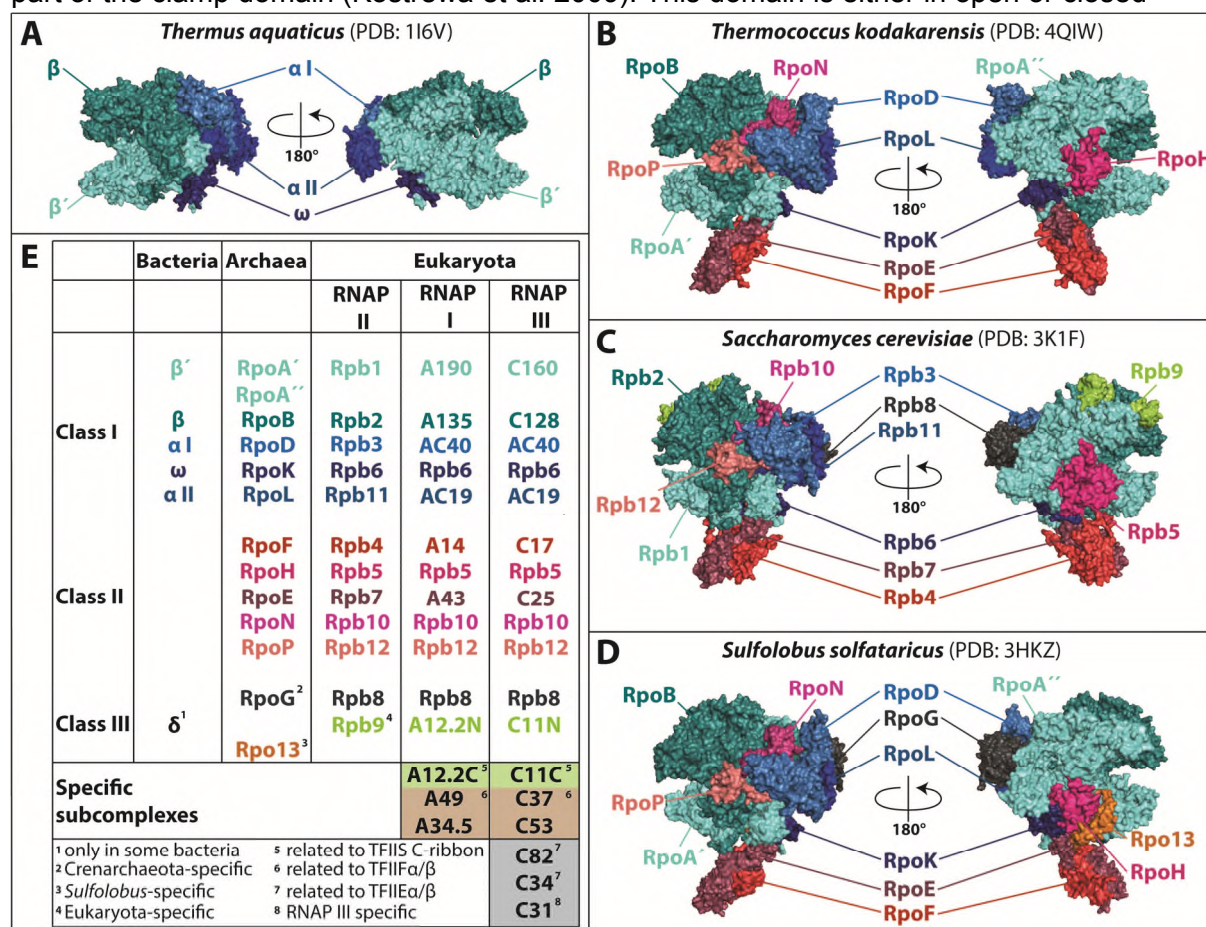
3.   The RNA polymerase

DNA-dependent RNA polymerases are the central enzymes for gene expression in all living cells. These enzymes belong to the conserved protein family of multi-subunit RNA polymerases, and possess a conserved core consisting of five subunits within all three domains of live (Werner 2008) (Figure 4). The overall structure of the core polymerase looks like a `crab claw`, and the two largest subunits form the pincers which defines the main cleft with a 25Å width for DNA loading (Ebright 2000) (Figure 5). Based on the size of the different subunits, the nomenclature is Rpb1 for the largest and Rpb12 for the smallest subunit in eukaryotes, and Rpo1 - Rpo13 in archaea, respectively, whereas latter are often referred to RpoA´ - RpoG (Figure 4). All subunits can be grouped into three classes. Class I contain the catalytic subunits, class II the subunits important for the assembly of the holo-enzyme, and subunits belonging to class III are auxiliary subunits (Werner, Grohmann 2011) (Figure 4). The assembly of this large multi-subunit enzyme occurs in a stepwise manner and was described for the eukaryotic Pol II enzyme (Wild, Cramer 2012). Three distinct assembling steps are forming the complete RNA polymerase, whereas the subunits Rpb10/12/11/3 form the first subcomplex which then interact with the Rpb2/9 intermediate, and in the last step the third subcomplex, Rpb1/5/6/8 associates to the other subunits to form the ten subunit core of the RNA polymerase (Wild, Cramer 2012).

The first crystal structure of the bacterial RNA polymerase was solved in 1999 for *Thermus aquaticus* with a resolution of 3.3Å (Zhang et al. 1999), and for yeast RNAP II in 2000 with a resolution of 3.0Å, whereas this RNAP lacked the stalk subunits Rpb4/7 (Cramer et al. 2000). Eight years later, the complete RNAP crystal structure of the crenarchaeal organism *Sulfolobus solfataricus* was solved at a resolution of 3.4Å (Hirata et al. 2008), which completes the set of available crystal structures of RNA polymerase enzymes from all three domains. Intensive structural comparisons of the three enzymes revealed an overall conservation in sequence, subunit composition, structure, function and mechanism, and therefore it was pointed out that all RNA polymerases originated from one latest universal common ancestor (LUCA), and the primordial RNAP resembles rather the bacterial RNAP structure (Korkhin et al. 2009; Werner, Grohmann 2011). The structural conservation was also confirmed by the exchange of subunits between Pol II and archaeal RNAP, e.g. RpoP, which is essential for growth, can be incorporated into a ΔRpb12 RNAP II and can complement its function *in vivo* and *in vitro* (Reich et al. 2009). In turn, Rpb12, which contacts Rpb2 and Rpb3, can replace RpoP in the archaeal RNA polymerase (Reich et al. 2009). A further exchange was performed with the eukaryotic Rpb5 subunit and the archaeal counterpart RpoH (Grünberg et al. 2010). RpoH is required for early steps of transcription initiation, and the activity of ΔRpoH RNAP can be rescued by Rpb5 (Grünberg et al. 2010). A chimeric Rpb5/RpoH construct consisting of the N-terminal yeast Rpb5 domain and the C-terminal RpoH of *Pyrococcus furiosus* RNAP, complement growth deficiencies of a ΔRpb5 RNAP II enzyme, but only if the corresponding amino acid E62 of RpoH is exchanged with lysine (Sommer et al. 2014).

The most conserved regions of the polymerase enzyme are located around the active center of the polymerase (Figure 5) on two beta-psi-barrels, forming the site for nucleotide incorporation (Ruprich-Robert, Thuriaux 2010). It is defined by the two largest catalytic subunits and each of these subunits are encoded by one gene in bacteria and eukaryotes, whereas the largest subunit is usually split into two genes RpoA´, and RpoA´´ in archaea (Werner 2007). The largest subunit comprises the highly conserved elements `lid´, `clamp coiled-coil domain´, `rudder´, `bridge´, and the `trigger loop´, a couple of `switches´ and a `funnel´, whereas the `fork loop 1´, `fork loop 2´and the `wall´ together with the `lobe´,

`protrusion´, and `switch´ are located at the second largest subunit (Ruprich-Robert, Thuriaux 2010). The clamp coiled-coil (CCC) domain is an important binding site for region 2 of the σ^{70} and TFB/TFIIB, respectively, and together with TFIIB it supports DNA melting and is further part of the clamp domain (Kostrewa et al. 2009). This domain is either in open or closed

Figure 4: Structural comparison of RNA polymerases of the three domains and their respective subunits. **A)** Crystal structure of the bacterial *Thermus aquaticus* RNAP (Zhang et al. 1999)(PDB: 1I6V), **B)** the euryarchaeal *Thermococcus kodakarensis* RNAP (Jun et al. 2014)(PDB: 4QIW), **C)** the eukaryotic *Saccharomyces cerevisiae* RNAP II (Kostrewa et al. 2009)(PDB: 3K1F) and **D)** the crenarchaeal *Sulfolobus solfataricus* RNAP (Hirata et al. 2008)(PDB: 3HKZ) are shown. Corresponding subunits are equally colored for each structure, whereas the following general color code is used: class I subunits forming the core are colored with different blue, class II subunits have red colors and are shared between archaea and eukaryotes, whereas class III subunits are unique colored and are not shared in every phylum. **E)** Overview of the subunits for bacterial, archaeal and eukaryotic Pol I - III enzymes Homologous subunits have the same color, and color code is the same as in A) - D)(Structure of Pol I and Pol III not shown. Taken from (Vannini, Cramer 2012)).

Table E content:

	Bacteria	Archaea	Eukaryota RNAP II	Eukaryota RNAP I	Eukaryota RNAP III
Class I	β′	RpoA′ RpoA′′	Rpb1	A190	C160
	β	RpoB	Rpb2	A135	C128
	α I	RpoD	Rpb3	AC40	AC40
	ω	RpoK	Rpb6	Rpb6	Rpb6
	α II	RpoL	Rpb11	AC19	AC19
Class II		RpoF	Rpb4	A14	C17
		RpoH	Rpb5	Rpb5	Rpb5
		RpoE	Rpb7	A43	C25
		RpoN	Rpb10	Rpb10	Rpb10
		RpoP	Rpb12	Rpb12	Rpb12
Class III	δ[1]	RpoG[2]	Rpb8	Rpb8	Rpb8
			Rpb9[4]	A12.2N	C11N
		Rpo13[3]			
Specific subcomplexes				A12.2C[5]	C11C[5]
				A49[6]	C37[6]
				A34.5	C53
					C82[7]
					C34[7]
					C31[8]

[1] only in some bacteria
[2] Crenarchaeota-specific
[3] *Sulfolobus*-specific
[4] Eukaryota-specific
[5] related to TFIIS C-ribbon
[6] related to TFIIFα/β
[7] related to TFIIEα/β
[8] RNAP III specific

conformation, which strongly depends on the interaction with the transcription factor E and the elongation factor Spt4/5, whereas the open conformation enables a better loading of the DNA, and the closed conformation increases the processivity of the RNAP due to function as a clamp to keep DNA in position (Grohmann et al. 2011; Schulz et al. 2016).

The rudder and the lid are both part of the largest subunit of the RNA polymerase and are located at the upstream edge of the RNA-DNA hybrid at the point where the t-strand of the transcription bubble reanneals with the non-template strand of the DNA (Kuznedelov et al. 2002; Naryshkina et al. 2006). Both are involved in the maintenance of the transcription bubble during elongation (Vassylyev et al. 2007), whereas the lid element supports the separation of RNA from DNA by interaction with nascent RNA at the upstream edge of the

RNA-DNA hybrid in elongation complexes (Naryshkina et al. 2006). The rudder element was suggested to stabilize the elongation complex via direct interactions with the nascent RNA (Kuznedelov et al. 2002) and is required for open complex formation (Naji et al. 2007). The highly flexible trigger loop discriminates between dNTPs and NTPs by direct interaction with the incoming substrate, and further plays a key role in transcription fidelity (Fouqueau et al. 2013). The bridge helix is also a flexible region which was thought to translocate the DNA after NTP incorporation, possibly by a kinking mechanism (Weinzierl 2011). Weinzierl showed that a small protein region termed the link domain is also involved in this process as this domain functions as a conformational sensor that recognizes pyrophosphate (PPi) and induces kinking of the bridge helix. Both, the TL and the bridge act in a coordinated manner (Weinzierl 2011). The fork loop 1 is part of the second largest subunit, and is absent in bacterial RNA polymerases (Gnatt et al. 2001). This region participate in the maintenance of the transcription bubble in elongation complexes, and was shown to be essential for transcription but not involved in preinitiation complex assembly in human (Jeronimo et al. 2004). In contrast, deletion of the fork loop 1 showed no requirement of this domain in promoter-dependent transcription in archaea (Naji et al. 2007). The fork loop 2 is part of the second largest subunit and is a small flexible region which was shown to
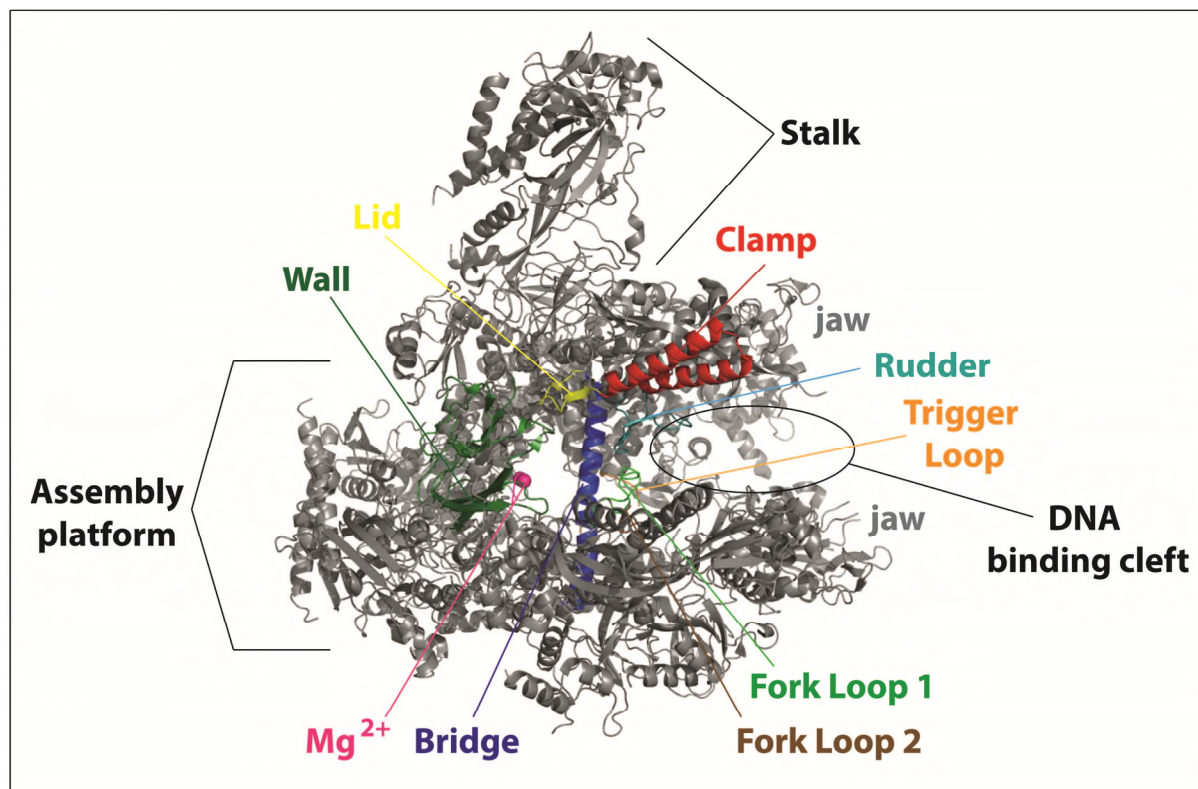
Figure 5: Crab claw structure and conserved regions of the RNA polymerase enzyme. The polymerase enzyme resembles a crab claw structure defined by the upper and lower jaw (grey), which forms the DNA binding cleft (indicated by a black ellipse). Conserved regions are highlighted using the following color code: Wall (dark green), Lid (yellow), clamp domain (red), rudder (cyan), trigger loop (orange), fork loop 1 (green), fork loop 2 (brown), bridge helix (blue), Mg^{2+}-ions (magenta). The active site for NTP incorporation is centered next to the Mg^{2+} ions and the bridge. Trigger loop and fork loop 2 are truncated in this structure due to the high flexibility. Lobe, protrusion, and switch regions are not highlighted. Structure modified from PDB: 4BBS (Sainsbury et al. 2013). Structural information taken from (Kostrewa et al. 2009; Naji et al. 2007; Weinzierl 2011; Fouqueau et al. 2013; Garcia-Lopez, Navarro 2011; Ruprich-Robert, Thuriaux 2010).

interact directly with the unpaired DNA at the nt-strand position i+2, which is two nucleotides next of the incoming substrate binding site (Kireeva et al. 2011). The fork loop 2 is further not involved in processes like DNA melting, translocation, and transcriptional activity, but it modulates interactions with the incoming NTP substrate and therefore regulates the catalytic step during transcription cycle (Kireeva et al. 2011). The highly conserved wall element is formed by the Rpb2 subunit, and this element enables binding of TFB/TFIIB with its B-core element, whereas a slight rotation was predicted together with the lobe element and the protrusion to partially close the cleft (Kostrewa et al. 2009; Sainsbury et al. 2013). The Mg^{2+} ions are located at the active site and are trapped by aspartate residues, whereas the first Mg^{2+} is permanently located at the incorporation site, and the second Mg^{2+} ion is recruited to interact with the phosphate residue of the incoming NTP (Sosunov et al. 2003). All RNAPs share the same mechanism of RNA synthesis and this process is termed the nucleotide addition cycle (Brueckner et al. 2009).

Despite the high level of conservation between the domains a few differences in structure exist. The most pronounced difference between the archaeal/eukaryotic RNA-polymerase and bacterial RNA polymerase is the stalk, which is defined by the subunits RpoE´ and RpoF in archaea, and their eukaryotic counterparts Rpb4 and Rpb7, respectively (Werner, Grohmann 2011). In yeast, Rpb4/7 interacts reversibly with the RNAP (Orlicky et al. 2001), and in archaea the stalk is stably incorporated (Grohmann et al. 2009). It was shown for the 11-subunit RNAP of *P. furiosus* that the core enzyme without the stalk is able to open DNA, but not to the full transcription bubble length, whereas addition of RpoE´ stimulates the transcription bubble formation, and RpoF does not influence the activity of RNAP during transcription (Naji et al. 2007). The stalk further increase the processivity rate of the RNAP during elongation (Hirtreiter et al. 2010b), and it was shown that the stalk is necessary for TFE activity due to the close location of TFE to the stalk, which induces a conformational change of the clamp to an open configuration to enable DNA loading and better DNA melting (Schulz et al. 2016).

A striking difference between archaeal and eukaryotic enzymes is the carboxy terminal domain of the largest subunit Rpb1 in eukaryotes, which lacks in archaeal organisms, and plays an important role in post-translational modifications during the transcription cycle to modulate activities in transcription and processing of nascent RNA (Hsin, Manley 2012). The subunit Rpb9 is a further example of differences between Archaea and Eukaryotes, as this subunit is exclusive for eukaryotic RNAPs, and it was shown that Rpb9 interact with the transcription factor IIF in eukaryotes (Ziegler et al. 2003). However, there is also a difference in the archaeal domain concerning the subunit composition. In Crenarchaeota, the subunit RpoG (Rpo8, respectively), a homolog of the eukaryotic Rpb8 is present, which lacks in the RNA polymerase of euryarchaeal organisms, and was first identified in the crystal structure of the *Sulfolobus shibatae* RNAP enzyme (Korkhin et al. 2009). This enzyme further contains a crenarchaeota-specific subunit Rpo13, which concrete function is still unknown, but it was hypothesized that it plays a role in initiation and elongation, due to the close location to the proximal path of the DNA (Wojtas et al. 2011).

In Eukaryotes, two further specific RNA polymerases emerged due to the increased complexity of regulation of transcription of specific genes, the RNA polymerase I and III. These enzymes contain 14 and 17 subunits, whereas the RNAP II contains 12 subunits, and ten subunits are shared within the three enzymes (Vannini, Cramer 2012). RNAP I contains the additional specific subunits A12.2C ribbon, and the subcomplex A49/A34.5 (Kuhn et al. 2007), the Pol III contains the C11C ribbon, the subcomplex C37/C53 and the subunits C82, C34 and the Pol III specific C31 (Wang, Roeder 1997; Landrieux et al. 2006). Structural

comparisons showed that the C11 C-ribbon and the A12.2 C-ribbon are related to the TFIIS C-terminus and therefore involved in RNA cleavage, whereas the C37/C53 subcomplex and the A49/A34.5 subcomplex are related to the TFIIFα/β subunits of the Pol II system and are important for initiation complex formation and start site selection, and the RNAP III specific subunit C31 stabilizes the open complex (Vannini, Cramer 2012). The C34 subunit functions in DNA opening (Brun et al. 1997), whereas the C82 subunit is related to the TFIIEα domain of the Pol II system (Wang, Roeder 1997).

Despite the degree of specialization of the eukaryotic RNA polymerases, the mechanism of RNA synthesis is highly conserved. In addition, the multi-subunit enzymes are platforms for transcription factors as a further point of transcription regulation.

4. The transcription factor E

TBP and TFB alone are sufficient to direct archaeal promoter-dependent transcription *in vitro* (Hausner et al. 1996; Qureshi et al. 1997; Bell, Jackson 1998; Soppa 1999). Eukaryotic RNAP II transcription can also be activated with TBP and TFIIB *in vitro*, using strong promoters and negatively supercoiled DNA templates (Parvin, Sharp 1993). However, a third factor associates to the core preinitiation complex (DNA-TBP-TFB-RNAP) which is termed transcription factor E (TFE). This factor is not strictly necessary for transcription initiation, but supports promoter opening and stabilization of the transcription bubble and the complex (Grünberg et al. 2007).

Eukaryotic TFIIE consists of TFIIEα and TFIIEβ, whereas most archaeal organisms express a single TFE, which is structurally similar and homolog to the TFIIEα subunit (Bell et al. 2001; Hanzelka et al. 2001). TFIIEα/TFE comprises a winged helix domain and a Zn-ribbon domain, whereas TFIIEβ contains two winged helix motifs important for protein-protein interactions and DNA binding (Tanaka et al. 2015). Point mutations of the human TFIIEα domain revealed that the N-terminal half of the winged helix motif binds to TFIIEβ helix-loop-helix domain, and the C-terminus of TFIIEα interacts with the N-terminal cyclin-fold of TFIIB (Tanaka et al. 2015). Interactions between TFIIEα and TFIIB, and between TFIIEβ and subunits p62 and p52 of TFIIH were demonstrated (Ohkuma et al. 1995), as well as contacts between TFIIEα and RNA polymerase subunits Rpb5, Rpb1 and Rpb2, and between TFIIEβ and RNA polymerase subunits Rpb2 and Rpb12 (Hayashi et al. 2005). In eukaryotes TFIIEα/β helps to recruit the transcription factor TFIIH, which is important for ATP-dependent DNA melting (Holstege et al. 1996), whereas deletion of TFIIEα results in impaired growth, suggesting an essential role *in vivo* (Kuldell, Buratowski 1997). For eukaryotic TFIIE a crosslink to the non-transcribed strand was observed in the open complex (Kim et al. 2000). In addition, eukaryotic RNAP III comprises two subunits C82 and C34 (human C62 and C39), which are homologous to TFIIEα and TFIIEβ, whereas RNAP I lack a related TFIIE subunit or protein (Carter, Drouin 2010; Vannini, Cramer 2012).

In archaea it was shown that TFE is not strictly required for transcription, but TFE stimulates transcription initiation at core promoters by enhancing DNA strand separation (Forget et al. 2004; Naji et al. 2007), at weak promoters (Hanzelka et al. 2001) and under TBP limiting conditions (Bell et al. 2001). Moreover, TFE can bind single stranded DNA nonspecifically, and using crosslinking experiments it was shown that the winged helix of TFE directly interacts with the non-transcribed strand of the transcription bubble at position -9 and -11 relative to the TSS (Grünberg et al. 2007). This finding suggested a role for TFE in DNA melting and stabilization of the transcription bubble, whereas the stimulatory effect of open complex formation was shown using $KMnO_4$ footprinting (Naji et al. 2007).

Cryo-electron microscopy experiments (He et al. 2013), as well as single molecule FRET experiments located the winged helix domain of TFIIEα in the preinitiation complex at the tip

of the clamp coiled coil domain of the RNA polymerase subunit 1, and the Zn-ribbon domain at the base of the RNAP clamp in close proximity to the subunits 4/7 of the RNA polymerase (Grohmann et al. 2011). In this study it was shown that the subunits Rpo4 and Rpo7 are necessary for TFE activity, and binding of TFE induces a conformational change in the clamp domain of the RNAP to an open configuration with a higher width of the DNA binding groove, which enables template DNA loading (Grohmann et al. 2011). Interestingly, TFE and TFB both interact with the clamp domain, which might be the reason, why TFE can rescue defect TFB mutations (Grünberg et al. 2007; Grohmann et al. 2011). The interaction of TFE with the RNAP and the opening of the flexible clamp domain were thought to be concomitant with DNA melting and template loading, whereas the interaction of TFE to the single stranded non-transcribed strand increases the stability of the transcription complex (Blombach et al. 2016). Studies in the yeast RNAP II system also demonstrated specific interactions of the TFIIEα WH domain and RNAP II clamp, and TFIIEβ tandem WH domain with DNA (Grünberg et al. 2012). Some members of the Crenarchaeota also encode the TFIIEβ subunit (Blombach et al. 2016). In *Sulfolobus acidocaldarius*, for example both TFIIE subunits are present, but TFIIEβ was shown to be essential (Blombach et al. 2015). In contrast to the eukaryotic and archaeal lineages, bacteria do not use TFE/TFIIEα-like proteins (Chakraborty et al. 2012), but interestingly, one factor, CarD shows similar functions in enhancing promoter opening, but this factor is not related (Bae et al. 2015; Davis et al. 2015).

5.   Additional eukaryotic transcription factors TFIIF and TFIIH

TFIIF is a heterodimeric protein consisting of two subunits, TFIIFα, or RNA polymerase associating protein (RAP) 74, and TFIIFβ or RAP30 (Flores et al. 1988), whereas the subunits in yeast termed Tfg1, Tfg2 and a third yeast-specific subunit Tfg3, which is not essential for transcription (Chafin et al. 1991). Structural analysis of human TFIIF showed that both subunits TFIIFα and TFIIFβ contain a winged helix domain, which in turn are connected to the dimerization module formed by the N-terminal regions of TFIIFα/β (Gaiser et al. 2000; Kamada et al. 2001). It was shown that TFIIF binds to RNAP II, whereas TFIIF is anchored via its dimerization module and transform the RNAP into a transcription initiation competent RNAP II/TFIIF complex (Burton et al. 1986; Eichner et al. 2010). Mobility shift assays with purified human RNAP II, TFIID, TFIIB, TFIIA and the TFIIE subunits revealed that the small subunit of TFIIF, TFIIFβ (RAP30) is sufficient for the recruitment of RNAP II to the promoter site (Flores et al. 1991). Beside the recruitment of RNAP II, a few functions for TFIIF were predicted in the initiation complex. Cryo electron microscopy of the human transcription initiation complex indicates that TFIIF stabilizes the downstream DNA of the cleft, and TFIIF seems to interact with the downstream BRE and with the B-core domain of TFIIB (He et al. 2013). TFIIF was shown to be involved in TSS selection (Ghazy et al. 2004), it stimulates first phosphodiester bond formation, and stabilize an early RNA-DNA hybrid (Khaperskyy et al. 2008). TFIIF further suppresses levels of abortive transcripts during initiation by increasing the processivity of the RNA polymerase, and together with TFIIH it supports promoter escape and prevent the arrest of RNAP II during early elongation (Yan et al. 1999). Paralogues of TFIIF are also present in RNAP I (A49-A34.5) and in RNAP III (C53-C37) (Geiger et al. 2010).

The last factor which completes the eukaryotic initiation complex is TFIIH. This factor is recruited by TFIIE, and is usually required for transcription *in vitro* (Maxon et al. 1994). TFIIH plays an important role in DNA melting and promoter escape (Goodrich, Tjian 1994; Holstege et al. 1996). TFIIH is a large multi-subunit complex consisting of 10 subunits in total, whereas six subunits XPD, p62, p52, p34, p8 and p44 forms the core module of the human TFIIE (Rad3, Tfb1, Tfb2, Tfb4, Tfb5 and Ssl1 in yeast, respectively), and three subunits CDK7,

cyclin H, and MAT1 form a kinase module (Kin28, Ccl1, and Tfb3 in yeast) (Gibbons et al. 2012).

The most important subunit for the function of TFIIH is XPB in human and Ssl2 in yeast, respectively. These subunits, together with XPD/Rad3 are ATPases, whereas both have 3´-5´ and 5´ - 3´ directionality (Schaeffer et al. 1993; Tirode et al. 1999). Cryo-EM analysis showed that TFIIH is located above the cleft of the RNA polymerase II and contact the PIC at two sites, one next to TFIIE, and one is located in proximity to the downstream DNA (He et al. 2013). It was suggested that TFIIH opens the DNA using ATP as energy source, but it was shown that opening of DNA can function independently of the Ssl2/XPB helicase activity (Lin et al. 2005). Moreover, these subunits act as a double stranded DNA translocase to coil the downstream non-template strand towards the RNAP II cleft, which creates torsion and unwinding of DNA, which supports DNA opening (Grünberg et al. 2012; Fishburn et al. 2015). In addition, TFIIH is involved in the DNA excision repair pathway via its XPD subunit, which opens the DNA to excise a mismatched nucleotide (Schaeffer et al. 1993; Egly, Coin 2011).

C.   From initiation to elongation and termination

After successful assembly of the complete preinitiation complex at the promoter site the double-stranded DNA has to be melted around the initially melted region containing the transcription start site (Pan, Greenblatt 1994; Hausner, Thomm 2001). The resulting single-stranded region is termed the transcription bubble and extends from -9 to +5 in *P. furiosus* (Spitalny, Thomm 2003), and has a size of 18 nucleotides in eukaryotes (Liu et al. 2010), and 14 base pairs in bacterial complexes (Hsu 2002). Recent studies showed that TBP and TFB are located closer to the surface of RNA polymerase in the archaeal system, which likely supports DNA melting via a reduced melting temperature (Nagy et al. 2015). In addition, the AT-composition of the initially melted region further reduces the energy required for separating both strands, whereas the resulting torsion of the DNA, which is induced by DNA bending and transcription factors, may also trigger DNA strand separation. However, in eukaryotes separation is basically ATP-dependent, and driven by the ATPase subunit of TFIIH (Holstege et al. 1996). Moreover, the clamp coiled-coil domain of the large subunit of the polymerase, together with the TFB/TFIIB B-linker domain support promoter opening, as these peptide regions keep the strands apart and stimulate DNA melting (Kostrewa et al. 2009). The transcribing strand slips into the cleft of the polymerase near the active site, which was thought to be stabilized by the TFB/TFIIB B-reader loop domain (Kostrewa et al. 2009; Sainsbury et al. 2013). The non-transcribed strand is further stabilized by direct interactions with TFE/TFIIE (Grünberg et al. 2007). The t-strand DNA is threaded through the active site and scanned for the correct initiation site by the help of the TFB/TFIIB B-reader helix (Giardina, Lis 1993; Kuehner, Brow 2006). The polymerase starts with RNA synthesis, leading to short 2-12 nucleotide RNAs in the so-called abortive initiation (Duchi et al. 2016). This process is not well understood, but it was hypothesized that the polymerase does not dissociate from the promoter and constantly re-start RNA synthesis (Carpousis, Gralla 1980). This process is repeated until a stable RNA-DNA hybrid is formed (Pal, Luse 2003). It was shown in the bacterial system that the production of aborted transcripts happens very fast until a 6-mer RNA is formed, then the initial complex undergoes pausing to overcome several checkpoints, and then the synthesis continues (Duchi et al. 2016). The t-strand is pulled through the active site of the polymerase during initial RNA synthesis, resulting in DNA scrunching, as the polymerase remains stationary through this event (Revyakin et al. 2006). When RNA reaches a length of six nucleotides, it was suggested that the B-reader loop

domain stimulates separation of the RNA strand from DNA due to charge dependent interactions (Sainsbury et al. 2013). It is further postulated that RNA clashes with the B-reader helix at a length of eight nucleotides (Kostrewa et al. 2009), whereas it is suggested that this interaction is necessary to guide the transcript towards the RNA exit channel of the polymerase (Sainsbury et al. 2013). It was further postulated that the B-reader is displaced during this process, as this domain was predicted to be in the path of the advancing 5´end of the RNA (Bushnell et al. 2004; Kostrewa et al. 2009; Sainsbury et al. 2013). The displacement of the corresponding domain was also shown for the bacterial σ-factor, as the 3.2 domain is in path of the 5´end of the RNA (Basu et al. 2014). This B-reader displacement may trigger the following collapse of the transcription bubble, which was hypothesized to take place at register +10/+11 (Spitalny, Thomm 2003; Luse 2013), and marks the end of promoter clearance in eukaryotic transcription systems (Pal et al. 2005). In the very last step of the initiation, it was suggested that the nascent RNA with a length of 12/13 nucleotides clashes with the TFB/TFIIB Zn-ribbon domain, which blocks the exit pore of the RNA polymerase, leading to a destabilization of TFB/TFIIB and causes a release of this transcription factor from the complex (Tran, Gralla 2008; Cabart et al. 2011; Sainsbury et al. 2013). TBP remains bound to the promoter (Xie, Reeve 2004b), and the RNA polymerase enters the elongation stage of transcription to continue synthesis of RNA.

Recent studies on the archaeal transcription system demonstrated that an elongation factor, Spt4/5, competes with the initiation factor TFE for binding to the polymerase during transition from initiation to elongation (Grohmann et al. 2011). Spt4/5 functions as a clamp, it binds across the main cleft above the DNA and closes the cleft like a clamp, which increases the processivity of the RNA polymerase during elongation (Schulz et al. 2016). The protein consists of two subunits, Spt4 and Spt5, which form a stable complex and stimulates transcription processivity *in vitro* (Hirtreiter et al. 2010a). The eukaryotic/archaeal subunit Spt5 possess a NGN sequence (NusG N-terminal), which has a high sequence identity to the monomeric bacterial elongation factor NusG, and is the important domain for RNAP interaction (Mooney et al. 2009). Archaeal and eukaryotic Spt5 bind to the small protein termed Spt4, which is a structural component to stabilize the Spt4/5 complex (Wenzel et al. 2010). The interaction with RNA polymerase can increase the processivity of the enzyme during elongation, whereas it was shown that a RNAP-Spt4/5 complex can inhibit recruitment to the promoter site, as the closure of the main cleft prevents DNA loading during the initiation step (Werner 2012). Archaeal/eukaryotic Spt5 and the bacterial NusG protein additionally contain KOW domains, which are platforms for other regulating proteins (Hirtreiter et al. 2010a). One factor which interacts with the NusG KOW domain is Rho, a termination factor in the bacterial transcription system (Pasman, Hippel 2000).

During elongation process a tight binding of a stable RNA/DNA hybrid is required (Nudler et al. 1997), but synthesis is often disrupted by pausing, arrest and stalling events due to missing or misincorporated substrates (Werner, Grohmann 2011). These events cause a reverse movement of the RNAP and is called backtracking (Erie 2002). This proofreading mechanism usually leads to a displacement of the RNA 3´end from the active site, and cleavage of RNA (Lisica et al. 2016). A specific cleavage <u>t</u>ranscription <u>f</u>actor (II)<u>S</u> (TFIIS) was identified in the archaeal and eukaryotic RNAP II system, which extends into the active site of the RNAP and supports cleavage of the 3´end of the RNA ((Izban, Luse 1992; Hausner et al. 2000; Kettenberger et al. 2004). Bacteria have universal proteins belonging to the TFS/TFIIS-related GreA family (Esyunina et al. 2016), and RNAP I and RNAP III contain TFIIS-related subunits A12.2 and C11 which possess own cleavage activity (Chedin et al. 1998; Kuhn et al. 2007). After release of the cut RNA it reanneals with the active site and

synthesis continues.

The transcription cycle ends with termination of transcription. Several models for the disassembly of the ternary elongation complex (TEC) were postulated for the respective RNAP system and are summarized in (Washburn, Gottesman 2015; Porrua et al. 2016). Termination of archaeal RNAP is almost similar to the termination mechanism described for RNAP III (Fouqueau et al. 2013) and is often controlled by U-stretches at the end of the RNA transcripts (Dieci et al. 2013). It was hypothesized that the stretch (six to seven in yeast, four to five in mammals (Orioli et al. 2011)), results in a weak dU:dA hybrid, which is poorly stable. This instable construct lead to backtracking and subsequent release of the RNAP III from the DNA template (Nielsen et al. 2013). The same mechanism was hypothesized for archaeal organisms. For *Thermococcus kodakarensis in vivo* (Santangelo et al. 2009), histone genes in *Pyrococcus furiosus* (Spitalny, Thomm 2008), as well as for *Methanothermobacter thermoautotrophicus* (Santangelo, Reeve 2006) it was shown that termination in archaea also depends on the presence of oligo(dT) sequences. With termination of transcription, the RNA transcript is released, and the RNA polymerase can reassociate to the promoter for a new initiation event to start the next transcription round.

D.   The replication protein A of *P. furiosus*

At the beginning of the thesis experiments were performed in our laboratory in collaboration with a French institute using the replication protein A of *P. abyssi* (*Pab*RPA). In network interaction studies a specific interaction of *Pab*RPA and RNAP was demonstrated (Pluchon et al. 2013). Due to this interaction, the influence of *Pab*RPA on transcription was investigated, and it was shown that the presence of *Pab*RPA in the *P. furiosus* transcription system stimulates the formation of RNA transcripts *in vitro* (Pluchon et al. 2013), whereas the function of RPA in transcription is unknown. To analyze the possible role of the *P. furiosus* RPA (*Pfu*RPA) during transcription, several experiments were performed.

RPA is a member of the single stranded binding (SSB) protein family, which are distributed among the three domains, and are usually present at single stranded DNA, like replication bubbles, damaged DNA sites, and recombination intermediates (Fanning et al. 2006). Binding of these proteins to single stranded DNA prevents the reannealing and degradation of DNA and can help to recruit proteins involved in DNA synthesis and repair (Fanning et al. 2006). RPA is a heterotrimeric complex consisting of three subunits RPA41, RPA14 and RPA32 in *Pyrococcus* strains (Komori, Ishino 2001). The large archaeal subunit RPA41 shows sequence similarities to the eukaryotic RPA70 subunit, but the two corresponding smaller subunits have no sequences in common (Komori, Ishino 2001). Sequence alignments of euryarchaeal RPAs suggest a closer relationship to eukaryotic RPAs than to the SSBs in Crenarchaeota or bacterial SSBs (Pluchon et al. 2013), whereas the crenarchaeal SSB is striking similar to bacterial SSBs (Kerr et al. 2003). On the other hand, structural comparisons of the important oligonucleotide-oligosaccharide binding (OB) fold of crenarchaeal SSB indicate a relation to eukaryotic and euryarchaeal OB folds of RPA (Murzin 1993). The eukaryotic RPA70, as well as the *Pyrococcus* RPA41 comprises an additional Zn-finger motif which is strongly conserved (Bochkareva et al. 2000), whereas RPA70 harbors four (OB) folds, and the smaller subunits contain usually one OB fold (Skowyra, MacNeill 2012). However, different organisms encode variations of SSBs which differ in the number of OB-folds, and the presence of a Zn-finger motif (Skowyra, MacNeill 2012).

RPA of *Pyrococcus furiosus* interact with the recombination proteins RadA and Hjc as well as with replication proteins like the DNA polymerase, primase, proliferating cell nuclear antigen

(PCNA) and the replication factor C in immunoprecipitation assays (Komori, Ishino 2001). In *in vitro* experiments a stimulation of the RadA promoted strand exchange in presence of RPA was observed (Komori, Ishino 2001). In *Haloferax volcanii* the large subunit RpaC contains three OB folds which are essential for growth, whereas the two smaller subunits RpaA and RpaB are not (Skowyra, MacNeill 2012). In growth experiments in the *Thermococcus kodakarensis* ΔTrmBL2 strain, which lacks the DNA binding protein TrmBL2, RPA is upregulated possibly due to the DNA maintenance role of RPA (Maruyama et al. 2011).

Beside the role in replication and DNA maintenance it was shown that *Sulfolobus solfataricus* SSB stimulate transcription under TBP limiting conditions and in presence of the chromatin protein Alba, suggesting a role in chromatin disruption at the promoter and RNA polymerase recruitment (Richard et al. 2004). The interaction of RPA with the RNA polymerase was described for *Pyrococcus furiosus* in co-purification experiments (Komori, Ishino 2001), and further analysis with a quantitative proteomic approach also showed that RPA is collocated with RNAP II on transcribed regions of active genes (Sikorski et al. 2011). In high throughput protein interaction studies yeast Rfa1 (=RPA70, RPA41) is associated with RNAP II (Krogan et al. 2006). This finding leads to the assumption that RPA is also part of the elongation complex and binds to and stabilize the transcription bubble (Sikorski et al. 2011). However, the exact role of the replication protein A in transcription and the molecular interactions with the transcription machinery are still unknown.

E.   Scientific questioning of this thesis

At the beginning of this thesis a crystal structure of a eukaryotic initially transcribing complex (ITC) was solved at a resolution of 3.4Å (Sainsbury et al. 2013). This structure consists of RNAP II, TFIIB, and a 6nt RNA, which is bound to the transcribed strand (Figure 6). It was demonstrated that the TFIIB B-reader domain, consisting of the B-reader helix and loop region, is in close proximity to the active site of the polymerase. Several aspects with respect to the mechanism of transcription initiation were postulated based on this structure. First, it was assumed that the B-reader loop stabilizes the initial transcription bubble due to the close location to the t-strand. Second, the B-reader loop interact with the nascent RNA with a length of 6nt. Based on the negative charge of the yeast TFIIB B-reader, it was suggested that RNA is separated from DNA via charge-dependent interactions between the loop region and the 5´end of the nascent RNA. Third, the advancing RNA clashes with the B-reader helix to be guided towards the exit channel of the polymerase after it is separated from DNA. Fourth, because the Zn-ribbon blocks the exit pore of the RNAP II, TFIIB has to be released from the transcription complex to enter productive elongation, whereas this release is possibly triggered by a clash of the RNA with the Zn-ribbon at a length of 12/13nt. However, these postulations derived from structural analysis, which represents one image at a distinct point during transcription.

Structural information of the archaeal transcription initiation with respect to topological transitions in the initially transcribing complex and the transition to elongation lacks, especially for the transcription factor B of the organism *P. furiosus*.

To investigate dynamic processes of *Pyrococcus furiosus* TFB (*Pfu*TFB) during transcription initiation, an approach was used in this study which is based on UV-inducible unnatural amino acids and specifically radiolabeled DNA templates in the archaeal *Pyrococcus furiosus in vitro* transcription system. Distinct amino acids of the *Pfu*TFB, which correspond to amino acids of the related yeast TFIIB, were substituted with an unnatural amino acid, which reacts with C-H-bonds to form a covalent linkage when exposed to UV-light. In addition, the DNA is site-specifically radiolabeled at selected sites on the template. Both, the labeled DNA and the

respective TFB variants are used in *in vitro* transcription reactions together with TBP, RNA polymerase and TFE to map contacts between TFB and DNA. If the crosslinker is in proximity to the labeled site, the TFB is covalently bound to the DNA, whereas a radiolabeled band at the size of TFB can be monitored using SDS-PAGE and autoradiography. This method was basically established by Dr. Robert Reichelt in the laboratory of the associate Professor Michael Bartlett of the Portland State University, Ohio, USA. To investigate dynamic rearrangements of the transcription factor B, crosslinking experiments were performed in stalled complexes at distinct registers from initiation to early elongation. Using these experiments the movement of TFB domains during transcription initiation can be analyzed which in turn allows studying interactions between TFB and RNA/DNA, as well as topological transitions of the TFB within the transcribing complex.

Beside the transitions of TFB, the role of the B-reader loop during transcription initiation was investigated using an alanine screening approach of selected B-reader amino acids, which were thought to be involved in transcription bubble stabilization and RNA-DNA strand stabilization. Here, the TFB alanine variants were tested in several *in vitro* transcription assays and the impact of the respective mutation is analyzed.

*Pfu*TFB was further analyzed in single molecule FRET experiments to figure out its function in DNA bending, which takes place at the beginning of transcription initiation.

In addition, the replication protein A of *Pyrococcus furiosus* and its possible role in transcription was characterized using different *in vitro* techniques.
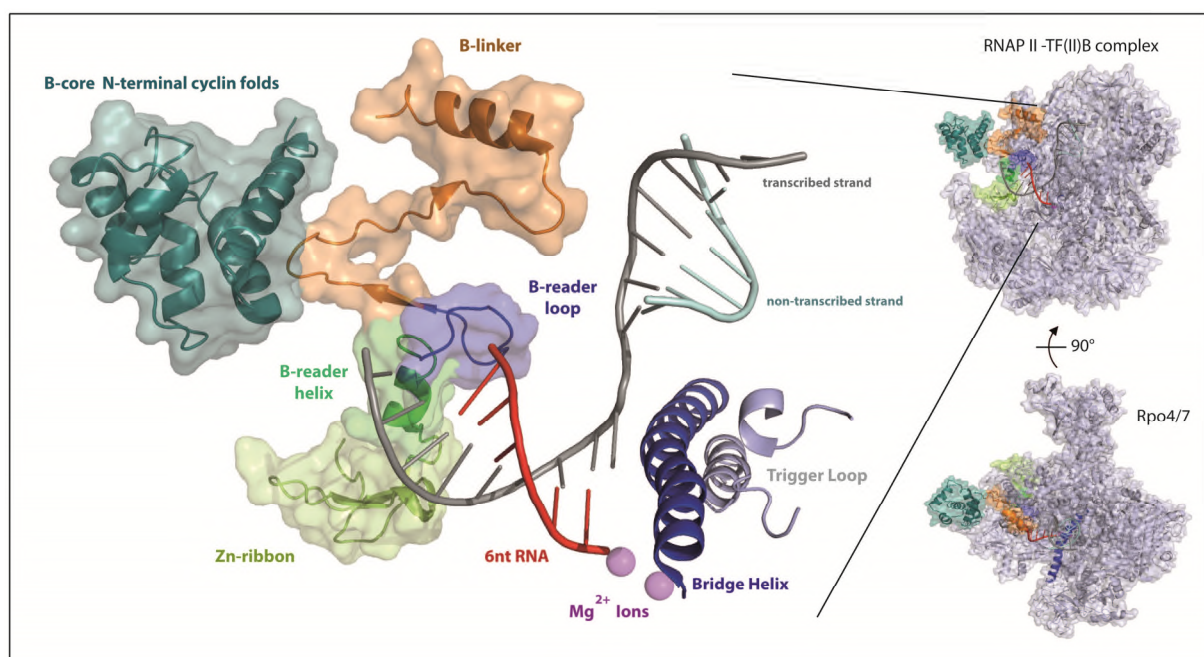
Figure 6: Crystal structure of the initially transcribing complex of *S. cerevisiae*. The structure contains the RNAP II (grey), TFIIB, the transcribing (dark grey) and non-transcribing strand (light blue) and an additionally 6nt RNA (red) which is bound to the t-strand. The center of the complex is enlarged to visualize the active site of the polymerase. This site is indicated by the Mg^{2+} ions (magenta), the bridge helix (dark blue) and the trigger loop (silver grey). The TFIIB consists of the N-terminal Zn-ribbon (yellow), the C-terminal B-core (cyan), the B-linker region (orange) and the B-reader, whereas this domain is separated into the B-reader helix (green) and the B-reader loop (blue). This B-reader domain is in close proximity to the active site of the polymerase. The C-terminal cyclin folds of the B-core domain and the connection between the B-linker and the B-core are missing due to insufficient resolution of the structure. The structure is modified from (Sainsbury et al. 2013).

II.   Materials

A.   Chemicals and Reagents

Substance	Supplier
Acrylamide, 2x	Serva, Heidelberg
Bromphenol blue	Serva, Heidelberg
dIC (poly 2´deoxyinosinic-2´-deoxycytidylic acid)	Sigma Aldrich, Steinheim
DTT (1,4 Dithiothreitol)	Serva, Heidelberg
dNTP Mix (100mM each)	Thermo Scientific, Waltham, USA
Dynabeads® His-Tag Isolation and Pulldown	Invitrogen (life technologies), Darmstadt
Dynabeads® M-280 Streptavidin	Invitrogen (life technologies), Darmstadt
Ethidium bromide	Serva, Heidelberg
Formamide	Merck, Darmstadt
GeneRuler 100bp / 1kb DNA ladder	Thermo Scientific, Waltham, USA
Glycogen (20mg/ml)	Roche, Mannheim
Heparin	Sigma Aldrich, Steinheim
IPTG (Isopropyl β-D-1-thiogalactopyranoside)	Roth, Karlsruhe
Lysozyme	Boehringer, Mannheim
Leupeptin	Sigma Aldrich, Steinheim
Mineral Oil	Roth, Karlsruhe
N,N´- Methylene bisacrylamide, 2x	Serva, Heidelberg
NTP´s (ATP, CTP, GTP, TTP)	Thermo Scientific, Waltham, USA
PageRuler Prestained Protein Ladder	Thermo Scientific, Waltham, USA
p-Benzoyl-L-phenylalanine	Bachem, Rubendorf, Switzerland
PCI (Phenol/Chloroform/Isoamyl alcohol)	Roth, Karlsruhe
Pepstatin A	Sigma Aldrich, Steinheim
Piperidine	Fluka, St. Gallen, Switzerland
PMSF (Phenylmethanesulfonyl fluoride)	Roth, Karlsruhe
Potassium permanganate	Merck, Darmstadt
Rotiphorese Gel 30	Roth, Karlsruhe
Scintillation Cocktail	Roth, Karlsruhe
TEMED (N, N, N´, N´-Tetramethylethylenediamine)	Roth, Karlsruhe
Tween 20	Thermo Scientific, Waltham, USA
Urea	Merck, Darmstadt
Xylene cyanol	Roth, Karlsruhe
[α-^{32}P] UTP, [α-^{32}P] dATP, [γ-^{32}P] ATP	Hartmann Analytic, Braunschweig
β-Mercaptoethanol	Roth, Karlsruhe

Standard chemicals and reagents which are not listed here were obtained from Merck (Darmstadt), Roth (Karlsruhe), Serva (Heidelberg), and VWR (Darmstadt).

B.   Kits

Kit	Supplier
DNA Cycle Sequencing Kit	Jena Bioscience, Jena
Wizard® SV Gel and Clean-Up System	Promega, Mannheim
Plasmid Miniprep Kit, peqGOLD	VWR, Darmstadt

C.   Enzymes

Enzyme	Supplier
BSA (20mg/ml, special quality)	Roche, Mannheim
Dnase I, RNase free	Thermo Scientific, Waltham, USA
EcoRI	Thermo Scientific, Waltham, USA
Klenow Fragment, exo⁻	Thermo Scientific, Waltham, USA
Lambda Exonuclease	New England Biolabs, Ipswich, USA
Lysozyme	Roth, Karlsruhe
Phusion HF DNA Polymerase	Thermo Scientific, Waltham, USA
S1 nuclease	Thermo Scientific, Waltham, USA
T4 DNA Ligase	Thermo Scientific, Waltham, USA
T4 Polynucleotide Kinase	New England Biolabs, Ipswich, USA

D.   Strains

Strain	Usage
E.coli BL21 (DE3) Star pEVOL Bpa	Expression of TFB-Bpa proteins
E.coli BL21 (DE3) pLysS	Expression of TFB-Ala and wtTFB proteins
E.coli BL21 (DE3) Codon Plus	Expression of TFE (*pf*0491)
E.coli DH5α	Selection/Storage of gdh-templates and mutated TFB sequences

E.   Services

Service	Supplier
Sequencing service	MWG eurofins, Ebersberg; Seqlab, Göttingen
Oligonucleotide synthesis	Biomers, Ulm; MWG eurofins, Ebersberg
Radioactive labeled nucleotides	Hartmann Analytic, Braunschweig

F.   Softwares

Software	Source (Reference)
Aida Image software v4.27	Raytest, Straubenhardt
BioEdit 7.2.5	http://www.mbio.ncsu.edu/bioedit/page2.html (T.A. Hall, 1999)
ImageJ	http://imagej.nih.gov/ij/(Abramoff, et. al., 2004)
ClustalOmega	http://www.clustal.org/omega/ (Sievers et al. 2011)
Pymol (V.0.99rc6)	https://www.pymol.org/ (Schrödinger, LLC)
iSMS 1.05	http://isms.au.dk (Preus et al. 2015)

G. Plasmids

Plasmid	Description	Source
pMUR125	pET14-TFB	(Goede 2004)
pMUR400	pET14-TFB-StopG41	This work
pMUR401	pET14-TFB-StopP42	This work
pMUR320	pET14-TFB-StopE43	This work
pMUR216	pET14-TFB-StopW44	(Zeller 2011)
pMUR217	pET14-TFB-StopR45	(Zeller 2011)
pMUR218	pET14-TFB-StopA46	(Zeller 2011)
pMUR219	pET14-TFB-StopF47	(Zeller 2011)
pMUR319	pET14-TFB-StopD48	This work
pMUR402	pET14-TFB-StopA49	This work
pMUR403	pET14-TFB-StopS50	This work
pMUR404	pET14-TFB-StopQ51	This work
pMUR220	pET14-TFB-StopR52	(Zeller 2011)
pMUR405	pET14-TFB-StopE53	This work
pMUR406	pET14-TFB-StopR54	This work
pMUR407	pET14-TFB-StopR55	This work
pMUR221	pET14-TFB-StopS56	(Zeller 2011)
pMUR222	pET14-TFB-StopR57	(Zeller 2011)
pMUR408	pET14-TFB-StopE74	This work
pMUR225	pET14-TFB-StopM85	(Zeller 2011)
pMUR409	pET14-TFB-StopF192	This work
pMUR232	pUC19gdhC6modCL2	Reichelt, R.
pMUR234	pUC19-gdhC5	(Spitalny, Thomm 2003)
pMUR235	pUC19-gdhC8	(Spitalny, Thomm 2003)
pMUR236	pUC19-gdhC9	(Spitalny, Thomm 2003)
pMUR237	pUC19-gdhC10	(Spitalny, Thomm 2003)
pMUR239	pUC19-gdhC15	(Spitalny, Thomm 2003)
pMUR8	pUC19-gdhC20	(Spitalny, Thomm 2003)
pMUR323	pUC19-gdhC45	Reichelt, R.
pMUR410	pUC19-gdhC11new	This work
pMUR411	pUC19-gdhC12new	This work
pMUR412	pUC19-gdhC13new	This work
pMUR413	pUC19-gdhC14new	This work
pMUR414	pUC19-gdhC15new	This work
pMUR177	pET14-TFE	(Zeller 2011)
pMUR419	pET14-TFB-R52A	Dick, C. (Bachelor thesis)
pMUR420	pET14-TFB-E53A	Dick, C. (Bachelor thesis)
pMUR421	pET14-TFB-R54A	Dick, C. (Bachelor thesis)
pMUR422	pET14-TFB-R55A	Dick, C. (Bachelor thesis)
pMUR423	pET14-TFB-R52E53A	Dick, C. (Bachelor thesis)
pMUR424	pET14-TFB-E53R54A	Dick, C. (Bachelor thesis)
pMUR425	pET14-TFB-R54R55A	Dick, C. (Bachelor thesis)
pMUR426	pET14-TFB-R52-R55A	Dick, C. (Bachelor thesis)

H.   Oligonucleotides

Oligonucleotides for TFB-Bpa mutagenesis:

Oligo name	Sequence (5´ → 3´ direction)
TFBStopG41F-P	5´ Phosphate - TAATTGATATGTAGCCTGAGTG -3´
TFBStopG41R-P	5´ Phosphate - TGTTCTCTTCTATTACATAACC -3´
TFBStopP42F-P	5´ Phosphate - ATATGGGTTAGGAGTGGC -3´
TFBStopP42R-P	5´ Phosphate - CAATTATGTTCTCTTCTATTACA -3´
TFBStopE43F-P	5´ Phosphate - GTCCTTAGTGGCGTGCT -3´
TFBStopE43R-P	5´ Phosphate - CCATATCAATTATGTTCTCTTCTAT -3´
TFBStopD48F-P	5´ Phosphate - GCCACTCAGGACCCATATC -3´
TFBStopD48R-P	5´ Phosphate - GATATGGGTCCTGAGTGGC -3´
TFBStopA49F-P	5´ Phosphate - GCTTTTGATTAGTCTCAA -3´
TFBStopA49R-P	5´ Phosphate - ACGCCACTCAGGAC -3´
TFBStopS50F-P	5´ Phosphate - GCTTAGCAAAGGGAACG -3´
TFBStopS50R-P	5´ Phosphate - ATCAAAAGCACGCCACT -3´
TFBStopQ51F-P	5´ Phosphate - GATGCTTCTTAGAGGGAACGCA -3´
TFBStopQ51R-P	5´ Phosphate - AAAAGCACGCCACTCAGGAC -3´
TFBStopE53F-P	5´ Phosphate - CTCAAAGGTAGCGCAGGTCT -3´
TFBStopE53R-P	5´ Phosphate - AAGCATCAAAAGCACGCC -3´
TFBStopR54F	5´- CAAAGGGAATAGAGGTCTAGAACTGG -3´
wtTFB1StopR54R	5´ Phosphate - AGAAAGCATCAAAAGCACGCC -3´
TFBStopR55F-P	5´ Phosphate - AAGGGAACGCTAGTCTAGAACT -3´
TFBStopR55R-P	5´ Phosphate - ATCAAAAGCACTGAGAAGC -3´
wtTFB1StopE74R	5´- AAAGCCCCTTGTCATGAAGAAG -3´
wtTFB1StopE74F2	5´- CAACTTAGATTGGAATTGACAGATCG -3´
TFBStopF192F-P	5´ Phosphate - AAGTTACAGATAGATTGCGAGAAA -3´
TFBStopF192R-P	5´ Phosphate - CTTCCAATTTCCTTTTTATCAACT -3´

Oligonucleotides for TFB-Ala substitutions

(The primers listed here were used in the Bachelor thesis of Christopher Dick):

Oligo name	Sequence (5´ → 3´ direction)
R52-Ala-F	5´ Phosphate- TCTCAAGCGGAACGCAGGT -3´
E53-Ala-F	5´ Phosphate- TCTCAAAGGGCTCGCAGGT -3´
R54-Ala-F	5´ Phosphate- TCTCAAAGGGAAGCGAGGTCTAGA -3´
R55-Ala-F	5´ Phosphate- TCTCAAAGGGAACGCGCATCTAGA -3´
R52E53 Ala F	5´ Phosphate- TCTCAAGCGGCACGCAGGT -3´
E53R54 Ala F	5´ Phosphate- TCTCAAAGGGCTGCGAGGTC -3´
R54R55 Ala F	5´ Phosphate- TCTCAAAGGGAAGCGGCGTCTAG -3´
Loop Ala REV	5´ - TTGAGAAGCATCAAAAGCACG -3´
Loop Ala FOR	5´ Phosphate- GCAGCCGCGGCATCTAGAACTGGTGCACCAGAA -3´
Loop-Ala-R	5´ Phosphate- AGCATCAAAAGCACGCCACT -3´

Oligonucleotides for gdh-C-template generation:

Oligo name	Sequence (5´ → 3´ direction)
C11newF	5´ - AGGTCATTTGGAGGATATGGG -3´
C12newF	5´ - AGGTACTTTGGAGGATATGGG -3´
C13newF	5´ - AGGTAACTTGGAGGATATGGG -3´
C14newF	5´ - AGGTAATCTGGAGGATATGGG -3´
C15newF	5´ - AGGTAATTCGGAGGATATGGG -3´
C16-C19 Rev	5´ Phosphate - CATTAACGATACATTTTTGGGC -3´

Oligonucleotides for radioactive labeling of gdh-C-cassettes:

Oligo name	Sequence (5´ → 3´ direction)
gdh-C6modCL2 -4T	5´ - CCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCAAT TGGCATAACGAT -3´
gdh-C6modCL2-11T	5´ - CCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCAAT TGGCATAACGATACATGTT -3´
gdh-C6modCL2-19T	5´ - CCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCAAT TGGCATAACGATACATGTTAGGGCAAT -3´
gdh-C8-4T	5´ - CCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCCAT ATGCCATTAACGAT -3´
gdh-C8-19T	5´ - CCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCCAT ATGCCATTAACGATACATTTTTGGGCAAT -3´
gdh-C9-4tnew	5´ - CCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCCAT ATGCTCATTAACGAT -3´
gdh-C10-4t	5´ - CCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCCCT TGGCCTCATTAACGAT -3´
gdh-C10-19t	5´ - CCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCCCT TGGCCTCATTAACGATACATTTTTGGGCAAT -3´
gdh-C10-8nt	5´ - TAGATTCTTTGAGCCTAATCAAATAAACAAAAGGATTTCCACT CTTGTTTACCGAAAGCTTTATATAGGCTATTGCCCAAA -3´
-19t -C11 new	5´ - CCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCCAT ATCCTCCAAATGACCTCATTAACGATACATTTTTGGGCAAT -3´
-19t -C12 new	5´ - CCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCCAT ATCCTCCAAAGTACCTCATTAACGATACATTTTTGGGCAAT -3´
-19t -C13 new	5´ - CCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCCAT ATCCTCCAAGTTACCTCATTAACGATACATTTTTGGGCAAT -3´
-19t -C14 new	5´ - CCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCCAT ATCCTCCAGATTACCTCATTAACGATACATTTTTGGGCAAT -3´
gdh-C15-4tnew	5´ - CCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCCAT ATCCTCCGAATTACCTCATTAACGAT -3´
-19t -C15 new	5´ - CCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCCAT ATCCTCCGAATTACCTCATTAACGATACATTTTTGGGCAAT -3´
-19t -C20 new	5´ - CCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCCCT TGGCTCCAAATTACCTCATTAACGATACATTTTTGGGCAAT -3´

Oligonucleotides for template preparation and sequencing:

Oligo name	Sequence (5´ → 3´ direction)
M13FOR	5´- GCCAGGGTTTTCCCAGTCACGA -3´
M13REV	5´- GAGCGGATAACAATTTCACACAG -3´
M13FOR-Bio	5´ Biotin - GCCAGGGTTTTCCCAGTCACGA -3´
M13REV-Bio	5´ Biotin - GAGCGGATAACAATTTCACACAG -3´
M13-FOR-P	5´ Phosphate - GCCAGGGTTTTCCCAGTCACG -3´
M13-REV-P	5´ Phosphate - AGCGGATAACAATTTCACACAGG -3´
M13FOR-FAM	5´ (5)6-Carboxyfluorescein - GCCAGGGTTTTCCCAGTCACGA -3´
M13FOR-Cy5	5´ Cyanine 5 - GCCAGGGTTTTCCCAGTCACGA -3´
M13FCy3	5´ Cyanine 3 - GCCAGGGTTTTCCCAGTCACGA -3´
Mat2	5´- CCAAGCTTGCATGCCTGCAGGTCG -3´
T7FOR	5´- TAATACGACTCACTATAGGG -3´
T7REV	5´- CTAGTTATTGCTCAGCGGTG -3´
gdhShortFOR	5´- TAGATTCTTTGAGCCTAATCAAAT -3´
M13R-250	5´- TCATTAATGCAGCTGGCACGAC -3´
M13R-500	5´- TCAGGGGGGCGGAGCCTATG -3´
MiniStartF	5´- TTATATAGGCTATTGCCCAAAAATGTATTACTAATGAGGTAAT TTGGAGCATATGGGG -3´
MiniStartR	5´- CCCCATATGCTCCAAATTACCTCATTAGTAATACATTTTTGGG CAATAGCCTATATAA -3´

Miscellaneous oligonucleotides:

Oligo Name	Sequence (5´ → 3´ direction)
PfuC14-t-strand	5´ Cyanine 3 - CTAGAGGATCCCCCATATCCTCCAGATTACCTCATT AACGATACATTTTTGGGCAATAGCCTATATAAAGCTTTCGGTAAACAAGA GTGGAAATCCTTTT -3´
Gdh-C14-34t-nolabl	5´ - AAAAGGATTTCCACTCTTGTTTACCGAAAGCTTTATATAGGCT ATTGCCCAAAAATGTATCGTTAATGAGGTAATCTGGAGGATATGGGGGAT CCCTAG -3´
ssgdhCy5	5´ Cyanine 5 – ATTTCCACTCTTGTTTACCGAAAGCTTTATATAGGC TATTGCCCAAAAATGTATCGTTAATGAGGTAA -3´
SSV T6 nt-strand	5´ Biotin – CGGACCGAAGCGACCATCGCCGGAGAXTGGAGTAAAGTT TAAATACTG -3´ (X=T-Cy3b)
SSV T6 t-strand	5´ Atto647n – CAGTATTTAAACTTTACTCCAATCCCTGGCGATGGT CGCGCTTTCGGTCCG -3´

III. Methods

A. DNA preparations

DNA was amplified by PCR from pUC19 or pET14 plasmids using oligonucleotides, Phusion high fidelity DNA polymerase, the recommended buffer conditions and extension temperatures. Depending on the reaction, a Mastercycler personal (Eppendorf) or a Thermal Cycler DOPPIO (VWR) for a temperature gradient was used. All successful amplified DNAs were purified with Promega Wizard® SV Gel and Clean-Up System, eluted in TE´ buffer (10mM Tris pH 8.0, 0.1mM EDTA), and the final concentration was determined by photo spectrometry (NanoDrop ND1000). The quality and the specificity of the amplification were verified by 1% or 2% AGE.

1. DNA templates for *in vitro* transcription assays and EMSAs

If not otherwise noted, all PCRs were performed using 0.5µM primer with 60°C oligonucleotide annealing temperature for M13 primer pairs, 0.02U/µl Phusion DNA polymerase and 28pM plasmid DNA in 30 cycles. For *in vitro* transcription templates standard M13 primer pairs were used for the amplification of gdh-C cassettes on pUC19 plasmids. Fluorescently labeled M13 primers were used for EMSA templates, and biotinylated primers were used for the immobilization of the respective template DNA.

2. 5´end labeled templates for footprint experiments

A gdh-C20 DNA template was amplified from 28pM of the plasmid pMUR8 in 200µl PCR volume with 0.5µM of each biotinylated M13 reverse primer and M13 forward primer to immobilize the transcribed strand. The product was purified and the final concentration was determined. 1pmol DNA was attached to 10µg streptavidin-coated magnetic particles. Before nucleic acids were added to the magnetic particles, the beads were separated from supernatant with magnetic particle separator (MPS), and washed three times for 45 seconds with buffer A (10mM Tris/HCl pH 7.5, 100mM NaCl, 1mM EDTA). For 50pmol of DNA the beads were resuspended in 80µl buffer B (50mM Tris/HCl pH 7.5, 100mM NaCl, 1mM EDTA), and the template was added and incubated for 30 minutes under constant rotation at room temperature. The supernatant was discarded and the particles were dissolved in 80µl buffer C (10mM Tris/HCl pH 7.5, 1M NaCl, 1mM EDTA). The sample was left at room temperature for one minute, repeated one more time, and finally the beads were resuspended in TE´ buffer to an end concentration of 1pmol/µl. 29pmol of immobilized DNA were incubated with 40U PNK and 1.11MBq [γ^{32}P] ATP in a total volume of 40µl for 30 minutes at 37°C. After heat inactivation of the enzyme for 10 minutes at 60°C, the sample was placed on MPS, the supernatant was removed and the template was washed with 50µl H$_2$O, then DNA was dissolved in 29µl H$_2$O.

3. Mismatch template preparation

The protocol and the primer pairs were used and modified from (Zeller 2011). In independent reactions two templates were amplified from 28pM pMUR8 (pUC19 gdh-C20) with 0.5µM of each M13FOR and MiniStartR, and MiniStartF and M13REV, respectively. After purification and concentration determination 0.57nM of each amplified template were overlapped in a fusion PCR (A: 30sec 98°C, B: 5sec 98°C, C: 20sec 55°C, D: 20sec 72°C; repeat B-D x 8; E: 120sec 50°C, add primer pairs; F: 5sec 98°C, G: 30sec 60°C, H: 20sec 72°C; repeat F-H x 25; 5 minutes at 72°C) in two independent reactions with primer pair A (M13FOR-P/M13REV) to phosphorylate the nt-strand, and primer pair B (M13FOR/M13REV-P) to phosphorylate the t-strand. In addition, two PCRs with 28pM pMUR8 as template were
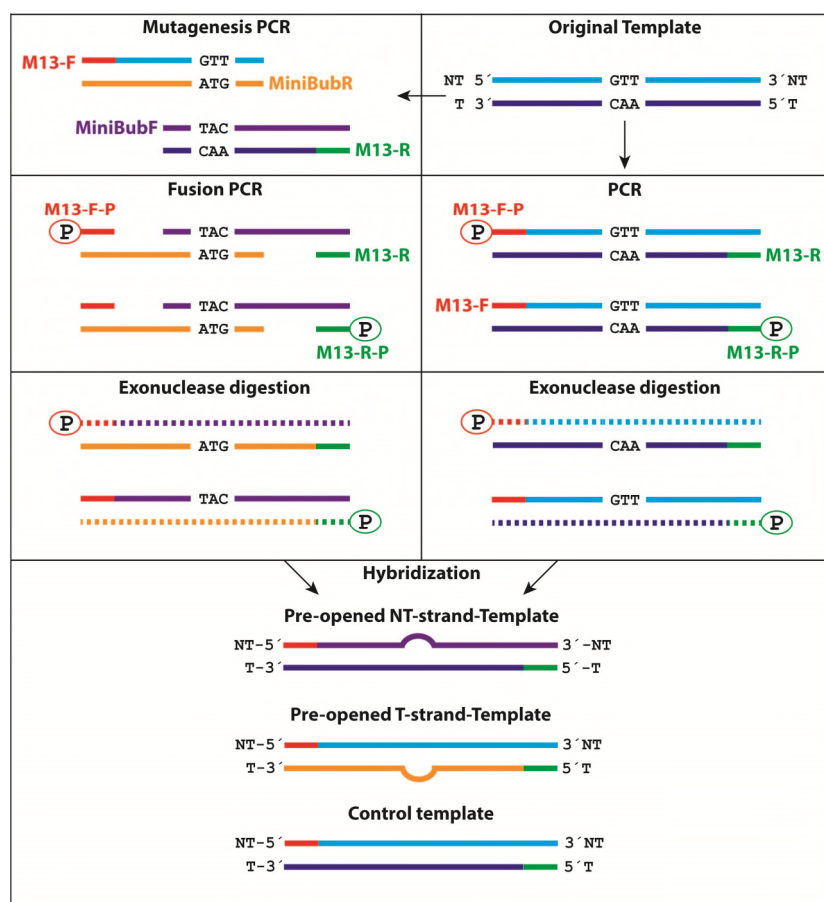
prepared with 0.5µM of each primer pair A and primer pair B. DNAs were purified and 4.5pmol of each were hydrolyzed with 10U λ-Exonuclease in a total volume of 50µl for 60 minutes at 37°C and 10 minutes at 75°C for inactivation of the nuclease. To receive an nt-strand mismatch, the nt-strand of the fused template was hybridized with the t-strand of the regular gdh-C20 cassette, and vice versa. Hybridization was performed at 95°C for 5 minutes, 45°C for 15 minutes and 40°C for 10 minutes. Quality of restriction and hybridization was verified by 2% AGE.

Figure 7: Schematic representation of pre-opened template preparation. The template is based on mutagenesis and fusion PCR of the original template, followed by Exonuclease digestion and hybridization. Modified from (Zeller 2011).
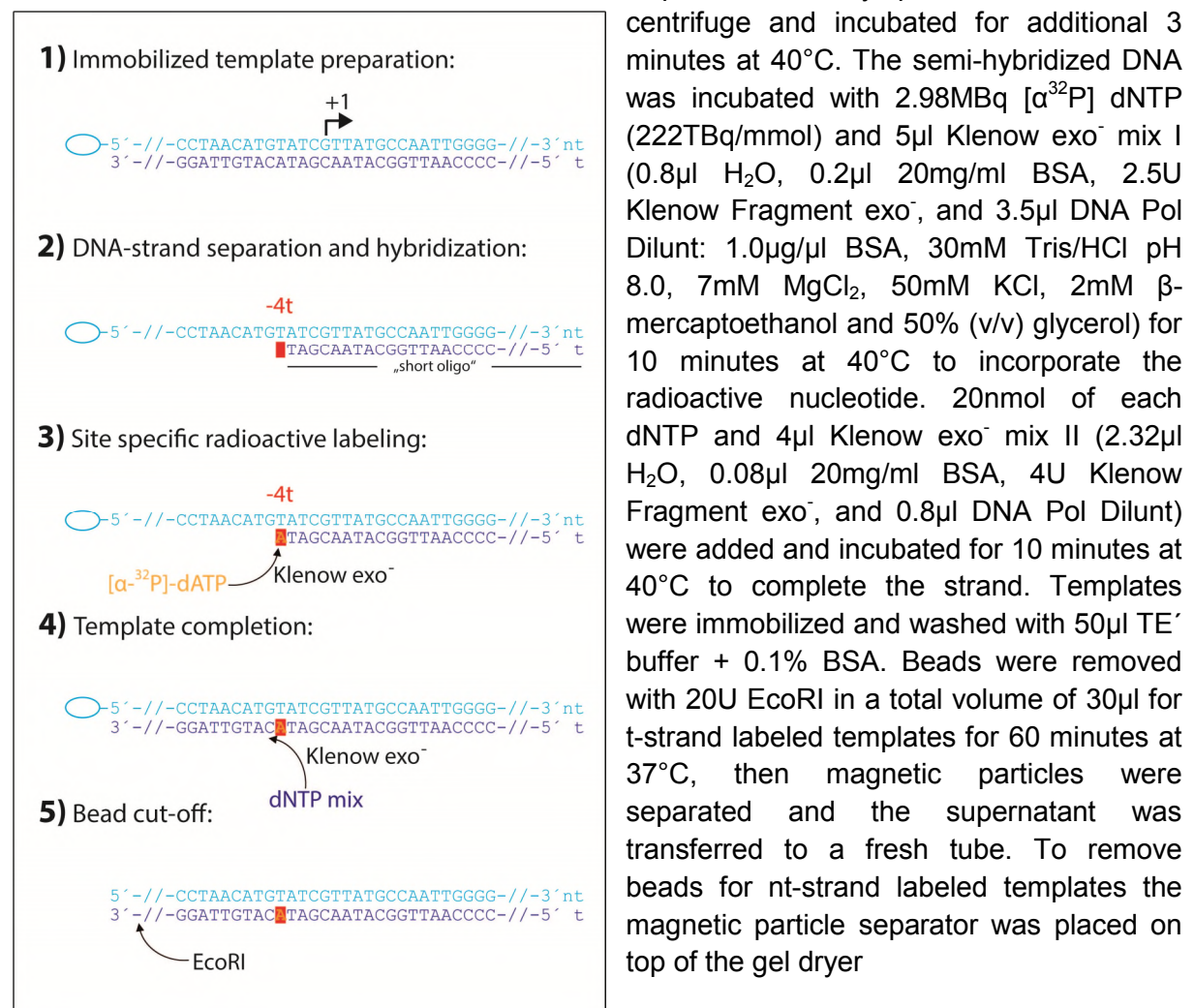
4. gdh-C11 - gdh-C15 template generation using PCR mutagenesis

The available plasmid pMUR8 (2.815bp, pUC19 containing the gdh-C20 cassette) served as basis for the generation of the new cassettes by PCR mutagenesis. The reactions were prepared in a total volume of 20µl with 0.55fmol pMUR8 and 0.5µM phosphorylated C11-C19REV primer in combination with 0.5µM of the respective forward primer C11newF - C19newF. These primers were designed to exchange the nucleic acid of interest via mismatch base pairing. A temperature gradient PCR with a range of 50°C – 70°C was used to identify the optimal annealing temperature of the primer pairs. After successful amplification the DNA was purified and 27fmol were re-circularized in a ligation reaction in a total volume of 50µl and 5U T4 DNA Ligase for 2 hours at 20°C. 2µl of the ligation sample was transformed using 50µl competent *E.coli* DH5α cells. First, the cells were defrosted on ice, and then the ligation product was added to the cells and left on ice for 15 minutes. The sample was heat-shocked for 40-60 seconds at 42°C and subsequently chilled on ice for additional 2 minutes. 350µl SOC medium (2% v/v Tryptone, 0.5% v/v yeast extract, 10mM NaCl, 2.5mM KCl, 10mM Glucose, 10mM $MgSO_4$, 10mM $MgCl_2$, pH 7.5) was added and the cells were incubated at 37°C for 60 minutes under constant rotation before plating on agar plates containing LB and 100µg/ml ampicillin. After growth two colonies were picked and each was inoculated in 5ml LB medium supplemented with 100µg/ml ampicillin. The cells were harvested and the plasmids were purified (Plasmid Miniprep Kit, peqGOLD) and sent to a sequencing service to verify the correct mutation. The success rate for mutagenesis was almost 100% for all cassettes, probably due to the usage of 5´phosphorylated primers, as it

is commonly used in QuickChange™ site directed mutagenesis protocol (Xia et al. 2015) or recommended in whole-plasmid mutations (Reikofski, Tao 1992).

5.   Radio labeled DNA templates for crosslink experiments

The protocol for radioactive labeling of double stranded DNA was used and modified from (Grünberg 2009). Templates for t-strand labeling were amplified from 28pM of the respective plasmid using 0.5µM of each Mat2 primer and biotinylated M13 forward primer. For labeling of the nt-strand 0.5µM of each gdhShortFOR for and biotinylated M13 reverse primer were used. The 216-231bp PCR products were purified and attached to streptavidin-coated magnetic particles to a concentration of 1pmol DNA per 1µl beads (Attaching protocol see III. A. 2). 10pmol of immobilized DNA was hybridized with 10pmol of the respective complementary single stranded oligonucleotide, while the 3´end of the oligo ends one nucleotide next to the position for labeling. Hybridization was done at 90°C for 3 minutes, 45°C for 15 minutes, and 40°C for 15 minutes. Samples were shortly spun down with a micro centrifuge and incubated for additional 3 minutes at 40°C. The semi-hybridized DNA was incubated with 2.98MBq [α^{32}P] dNTP (222TBq/mmol) and 5µl Klenow exo$^-$ mix I (0.8µl H$_2$O, 0.2µl 20mg/ml BSA, 2.5U Klenow Fragment exo$^-$, and 3.5µl DNA Pol Dilunt: 1.0µg/µl BSA, 30mM Tris/HCl pH 8.0, 7mM MgCl$_2$, 50mM KCl, 2mM β-mercaptoethanol and 50% (v/v) glycerol) for 10 minutes at 40°C to incorporate the radioactive nucleotide. 20nmol of each dNTP and 4µl Klenow exo$^-$ mix II (2.32µl H$_2$O, 0.08µl 20mg/ml BSA, 4U Klenow Fragment exo$^-$, and 0.8µl DNA Pol Dilunt) were added and incubated for 10 minutes at 40°C to complete the strand. Templates were immobilized and washed with 50µl TE´ buffer + 0.1% BSA. Beads were removed with 20U EcoRI in a total volume of 30µl for t-strand labeled templates for 60 minutes at 37°C, then magnetic particles were separated and the supernatant was transferred to a fresh tube. To remove beads for nt-strand labeled templates the magnetic particle separator was placed on top of the gel dryer

Figure 8: Schematic draw of the specific radioactive labeling of DNA templates. Immobilized dsDNA template is separated and hybridized using a short complementary single-stranded oligonucleotide. The [α^{32}P] dNTP is incorporated using Klenow exo$^-$ enzyme, and the strand is completed using unlabeled dNTPs. Beads were cut off with EcoRI in a last step.

(DrygelSr. Slab gel dryer, Hoefer scientific connected with HydroTech™ Vacuum Pump; Biorad) and heated up to 80°C, then samples were incubated for 5 minutes and supernatant was collected (Holmberg et al. 2005). Residual beads were washed with 50µl TE´ buffer and

after immobilization the supernatant was added to the collected sample. DNA was isolated with 85µl PCI and vigorous mixing for 40 seconds. After centrifugation at 21.000g for 5 minutes the upper phase was transferred to a new tube and the DNA was precipitated with ethanol using 1µl glycogen (20mg/ml), 30µl 3M sodium acetate pH 5.2, 185µl H_2O and 600µl ethanol (≥99.98% p.a.) for 45 minutes at -80°C. The sample was centrifuged at 21.000g for 30 minutes at room temperature, and then supernatant was discarded. The pellet was washed with 1ml 70% ethanol and centrifuged for 10 minutes at 21.000g. The supernatant was discarded, and the pellet was dehumidified using a vacuum concentrator (Concentrator 5301, Eppendorf). DNA was dissolved in 20µl TE´ buffer. The radioactive concentration was measured by placing 2x 1µl of the dissolved DNA on separate glass microfiber filter plates and dried until liquid was evaporated. 10ml scintillation cocktail together with one filter plate was placed in a scintillation tube and the activity was measured for 60 seconds in the Scintillation Analyzer (TriCarb 2900TR, Packard; QuantaSmart™ software). To define the concentration of the labeled DNA several dilutions of the [$\alpha^{32}P$] dNTP (222TBq/mmol) stock solution were analyzed with the Scintillation Analyzer to gain the relation between counts per minute (cpm) and the respective radioactive concentration. Using this method 1pmol of [$\alpha^{32}P$] dNTP (222TBq/mmol) was determined to have 1.52×10^7 cpm at the calculated date of activity.

B.   Protein preparations

The TATA binding protein (TBP) was cloned, expressed and purified from students during a molecular biology practical course and was originally described in (Hausner et al. 1996). This protein was diluted in 1xTB-0 transcription buffer (40mM Na-Hepes pH 7.3, 250mM NaCl, 2.5mM $MgCl_2$, 0.1mM EDTA, 0.1mM $ZnSO_4$, 10% v/v glycerol) to a final concentration of 50ng/µl and stored in small portions at -80°C.

The RpoD-his tagged RNAP was purified from *P. furiosus* cells, whereas the cells were kindly provided by the institute of Microbiology and Archaea Center, University of Regensburg.   .

The genetic sequence of the transcription factor B1 (TFB1; henceforth referred to as TFB) was fused to an C-terminal histidine tag and cloned into pET14 (pMUR125) by (Goede 2004). The protein was kindly provided by the institute, and later expressed and purified as described in (III. B. 3).

The replication protein A was kindly provided by Dr. Hausner, W., Institute of Microbiology and Archaea Center, University of Regensburg.

1.   SDS-PAGE and protein concentration

SDS-PAGE

Protein samples were dissolved in 6xSDS loading dye (70% v/v 0.5M Tris/HCl pH 6.8, 10mM SDS, 30% v/v glycerol, 10% w/v SDS, 0.6M DTT, 0.03% w/v bromophenol blue), and incubated at 95°C for 4 minutes and analyzed by 12% SDS-PAGE using the Biorad Mini-PROTEAN® Tetra Cell system, handcast gels (4% stacking gel: 2.4ml $H_2O_{millipore}$, 0.5ml Rotiphorese® 30 (37.5:1), 1.0ml 0.5M Tris/HCl pH 6.8 and 10mM SDS; 12% separating gel: 2.7ml $H_2O_{millipore}$, 3.0ml Rotiphorese® 30 (37.5:1), 1.9ml 1.5M Tris/HCl pH 8.8, 10mM SDS), running buffer (25mM Tris, 192mM glycine, 3.4mM SDS; (Laemmli 1970)) and the recommended manufacturer´s protocol (Biorad).

Protein concentration

The protein concentration was usually determined with Bradford method (Bradford 1976). In case of TFB the concentration was determined using 12% SDS-PAGE, whereas defined concentrations of the reference protein BSA and defined volumes of the purified TFB were separated. The gel was stained with Coomassie staining solution (40% v/v isopropanol, 10% v/v acetic acid, 0.2% w/v Coomassie R250) and destained with destaining solution (40% v/v isopropanol, 10% v/v acetic acid) to minimalize background. The gel was scanned and digitalized, and signals derived from proteins were quantified and compared with ImageJ software (Abramoff et. al., 2004).

2.   Mutagenesis of TFB

The site directed mutagenesis of TFB of *Pyrococcus furiosus* (pf1377) was achieved by using oligonucleotides which contained a 1-3nt mismatch to introduce a TAG sequence at distinct sites at the gene of TFB. For this reaction 14pM of the plasmid pMUR125 (5.566bp, pET14+TFB) was used together with 0.5µM of the respective oligonucleotides in a PCR with a total volume of 20µl. For optimal results a temperature gradient PCR was used to identify the best annealing temperature of the used primer pairs. Successful amplified templates were purified and 14fmol were re-circularized in a ligation step in a total volume of 50µl with 5U T4 DNA Ligase for 2 hours at 20°C. 2µl of the ligation sample were transformed in competent *E.coli* DH5α cells (For transformation see III. A. 4). Two colonies were picked from the plated transformation sample and inoculated in 5ml LB medium supplemented with 100µg/ml ampicillin and grown at 37°C. Plasmids were isolated (Plasmid Miniprep Kit, peqGOLD) and sent to a sequencing service to verify the successful nucleotide exchange of the wild type codon with the TAG-sequence at the expected site. The success rate was almost 100% for every TAG introduction, probably for similar reasons like for the template mutagenesis, chapter III. A. 4, namely the usage of phosphorylated primers (Reikofski, Tao 1992; Xia et al. 2015).

3.   Expression and purification of TFB and TFB variants

The plasmids containing the mutated TFB sequences were transformed in the expression strain *E.coli* BL21 (DE3) Star pEVOL Bpa (Young et al. 2010) for expression of TFB-Bpa, or in *E.coli* BL21 (DE3) pLysS for the expression of TFB-Ala and wild type TFB. Depending on the transformation efficiency of the expression strain 14 -28fmol plasmid DNA was added to 50µl of competent *E.coli* cells. The procedure for the heat-shock transformation is the same like for *E.coli* DH5α cells (see III. A. 4) whereas a 30 second incubations step at 42°C was used instead of the 40 - 60 seconds. Transformed cells were grown on agar plates with 100µg/ml ampicillin and 34µg/ml chloramphenicol. After incubation at 37°C one colony was picked and diluted in 200µl LB medium, and plated on an agar plate containing the respective antibiotics. After growth over night at 37°C the cells were removed from the plate using a spatula to inoculate 100ml of LB medium supplemented with 100µg/ml ampicillin and 34µg/ml chloramphenicol to an optical density at 600nm wavelength of 0.1 - 0.2 (Beckman DU®640 Spectrophotometer). Cells were grown at 37°C until OD_{600} reached 0.6 - 0.8. Cells were chilled on ice for five minutes and induced with 0.5mM IPTG for protein expression. For simultaneous incorporation of unnatural amino acids expression of modified tRNA´s encoded on the pEVOL plasmid was induced with 1.3mM L-arabinose in presence of 0.2mM *L*-p-benzoyl-phenylalanine (Bpa) (Kauer et al. 1986). Bpa was dissolved in 1M HCl, therefore an equal volume of 1M NaOH was used to adjust pH. Induced cells were incubated over night at 18°C at constant rotation in the dark (Infors HT Multitron Pro incubator). After induction the

cells were harvested by centrifugation at 4°C and 6.000g for 15 minutes (JA10 rotor, Beckman Coulter™ Avanti™ J-25). The pellet of a 100ml culture volume was dissolved in 3ml Binding/Wash buffer (50mM phosphate buffer pH 8.0, 300mM NaCl, 0.01% v/v Tween 20) and a small pile of Lysozyme was added and mixed. The suspension was incubated on ice for 45 minutes, and then the cells were treated with ultra sound for 6 x 1 minute (50% intensity, 50% pulse length; Branson sonifier 250) on ice and subsequently centrifuged at 4°C and 21.000g for 60 minutes (Hitachi himac CT15RE, VWR). The supernatant was transferred into a new tube and the samples were incubated at 70°C for 10 minutes to denature host proteins. The centrifugation step at 4°C and 21.000g for 60 minutes was repeated. For 100ml cell suspension 2 x 100µl His-Tag Dynabeads® were transferred in new tubes and placed at the magnetic particle separator to immobilize the particles. The supernatant was removed and the protein solution was added to the Dynabeads® and incubated for 30 minutes at 4°C under constant rotation (Rotamix RM1, Elmi). The magnetic particles were subsequently immobilized and supernatant was removed. The Dynabeads® were washed four times with 400µl Binding/Wash buffer by gentle pipetting for 45 seconds. Finally the protein was eluted by adding 100µl Elution buffer (50mM phosphate buffer pH 8.0, 300mM imidazole, 300mM NaCl, 0.01% Tween 20) to the Dynabeads® and incubated at 4°C for 30 minutes under constant rotation. The beads were immobilized and the eluted protein in the supernatant was stored at -80°C in small portions.

4. RNA-polymerase purification

2.0g of frozen *P. furiosus* cells (RpoD-strep-his cells) were defrosted on ice. Cells were dissolved in 5.0ml of Lysis buffer (100mM Tris/HCl pH 8.0, 1M NaCl, 20mM imidazole, 2.5mM $MgCl_2$, 20% v/v glycerol; protease inhibitor was added before use) until suspension turned into a homogenous solution. The sample was treated with ultra sound on ice for 6 x 1 minute (50% intensity, 50% pulse length, Branson sonifier 250). 2.5g glass beads (0.1mm diameter) were added and intensively mixed in a FastPrep®-24 (MP™) instrument for 4 x 30 seconds at 5.0M/s. The cell suspension was centrifuged at 4°C at 21.000g for 60 minutes (JA25.50 rotor, Beckman Coulter™ Avanti™ J-25). The supernatant was filtrated with a sterile syringe and a filter adapter (0.45µm pore diameter). The RNAP was separated with a HisTrap 1ml FF column at the ÄKTA purification system (Amersham Biosciences) and one-step eluted with Elution buffer (100mM Tris/HCl pH 8.0, 300mM NaCl, 250mM imidazole, 2.5mM $MgCl_2$, 20% v/v glycerol). The presence of the RNA-polymerase in different fractions was verified using a SDS-PAGE, and the respective samples were pooled and run on a Hiload™ 16/60 Superdex™ 75 size exclusion column with gel filtration buffer (100mM Tris/HCl pH 8.0, 150mM NaCl, 2.5mM $MgCl_2$, 20% v/v glycerol). The fractions containing the RNAP were collected and stored at -80°C in small portions. The stoichiometry of the RNA-polymerase subunits was verified by SDS-PAGE, the presence of subunits A´´, E´ and D was additionally verified using western blot experiments and antibodies against the respective subunits. Protein concentration was determined using Bradford analysis. Work solution was prepared by dissolving RNAP in gel filtration buffer to a final concentration of 100ng/µl and stored at +4°C.

5. TFE purification

The plasmid pMUR177 (pET14+TFE) (Grünberg 2009) was transformed into *E.coli* BL21 (DE3) Codon Plus (For transformation see III. A. 4) and grown on LB agar plate supplemented with 100µg/ml ampicillin and 34µg/ml chloramphenicol. One colony was inoculated in 200µl LB medium and plated on two separate LB agar plates supplemented with the respective antibiotics. After incubation over night at 37°C cells were collected using

a spatula and two flasks containing 400ml LB medium were inoculated to an OD_{600} of 0.1 (Beckman DU®640 Spectrophotometer). Cells were grown to an OD_{600} of 0.6 and then induced with 0.5mM IPTG for protein expression at room temperature over night. Cells were centrifuged at 6.000g for 15 minutes at 4°C (JA10 rotor, Beckman Coulter™ Avanti™ J-25) and supernatant was discarded. The pellets were dissolved in 5.0ml Lysis buffer (50mM Tris/HCl pH 8.0, 300mM NaCl, 10mM imidazole, 20% v/v glycerol), pooled, and a small pile of Lysozyme was added. After incubation at 4°C for 30 minutes the cell suspension was treated with ultra sound for 6 x 1 minute (50% intensity, 50% pulse length, Branson sonifier 250). A small pile of Dnase I was added and incubated for 10 minutes at 4°C. The suspension was centrifuged at 20.000g for 30 minutes at 4°C (JA25.50 rotor, Beckman Coulter™ Avanti™ J-25) and the supernatant was incubated at 70°C for 10 minutes. Centrifugation at 20.000g for 30 minutes at 4°C was repeated, and the supernatant was sterile filtered. The protein was separated with a HisTrap 1ml HP column and buffer A (20mM Na-Hepes pH 7.8, 500mM NaCl, 7mM $MgCl_2$, 10mM imidazole, 10% v/v glycerol) and one-step eluted with buffer B (20mM Na-Hepes pH 7.8, 500mM NaCl, 7mM $MgCl_2$, 0.01% v/v Tween 20, 200mM imidazole, 10% v/v glycerol) with an ÄKTA purification system (Amersham Biosciences). Fractions were analyzed with SDS-PAGE and samples containing the TFE protein were pooled and run on a HiLoad™ 16/60 Superdex™ 75 size exclusion column with gel filtration buffer (30mM Tris/HCl pH 7.5, 1mM EDTA, 300mM NaCl, 10% v/v glycerol). Important fractions were pooled and the protein was stored at -20°C in small portions.

C.   Transcription assays

Basically all samples were analyzed using 7M urea gels containing different polyacrylamide (PA) concentrations, except for EMSAs (native PAGE) and crosslink experiments (SDS-PAGE). The following urea-PA gels and electrophoretic conditions were used in this study: 7M urea 6% PA for $KMnO_4$ footprints at 2000V, 50W, 50mA for 3 - 4 hours (Biorad Sequi-Gen® Cell, gel size: 50cm x 21cm), 7M urea 8% PA for run-off transcription assays at 600V, 48W, 48mA for 3.5 hours (gel size: 30cm x 21cm), 7M urea 12% PA for the verification of incorporation of radiolabeled dNTPs into DNA at 400V, 40W, 25mA (gel size: 16cm x 13cm), and 7M urea 28% PA for abortive transcription assays at 2500V, 48W, 20mA for 12 hours (gel size: 40cm x 21cm). Finished gels were carefully transferred to a chromatography paper (3MM Whatman™, GE healthcare) and dried with a gel dryer (DrygelSr. Slab gel dryer, Hoefer scientific, connected to HydroTech™ Vacuum Pump; Biorad) for 35 - 50 minutes at 70°C. An imaging plate (FUJIFILM BAS-IP MS 2040) was placed on the gel for 30 minutes up to hours, depending on the used radioactive concentration. The IP was analyzed using Phosphoimager (FUJIFILM FLA 5000) and signals were detected and quantified with Aida v4.27 software.

1.   Electro mobility shift assay

Electro mobility shift assays were done using 26nM RNAP, 56nM 5´-FAM (6-Carboxyfluorescein)-labeled template DNA gdh-C20 or alternatively labeled templates, 475nM TBP and 270nM TFB in transcription buffer (40mM Hepes pH 7.3, 250mM NaCl, 2.5mM $MgCl_2$, 0.1mM EDTA, 0.1mM $ZnSO_4$), 1mM DTT and 0.1µg/µl BSA in a total volume of 10µl and incubated at 70°C for 10 minutes. Then 0.5µg dIC (poly-2´deoxyinosinic-2´-deoxycytidylic acid) was added and incubated for additional 10 minutes at 70°C. Samples were mixed with 2µl 85% glycerol and separated on handcasted native 4-10% or 4-20% gradient PAGE at 170V for 35 minutes (4% stacking gel: 2.4ml $H_2O_{millipore}$, 0.5ml Rotiphorese® 30 (37.5:1), 1.0ml 0.5M Tris/HCl pH 6.8; 4% separating gel: 1.545ml

H$_2$O$_{millipore}$，0.33ml Rotiphorese® 30 (37.5:1), 0.625ml 1.5M Tris/HCl pH8.8; 10% separating gel: 1.045ml H$_2$O$_{millipore}$, 0.83ml Rotiphorese® 30 (37.5:1), 0.625ml 1.5M Tris/HCl pH 8.8; 20% separating gel: 0.5ml H$_2$O, 1,375ml Rotiphorese® 40, 0.625ml 1.5M Tris/HCl pH 8.8; running conditions: 25mM Tris, 192mM glycine; Biorad Mini-PROTEAN® Tetra Cell system). The gel was analyzed using Phosphoimager (FUJIFILM FLA 5000) and Aida v4.27 software.

2.   Abortive transcription assay

Transcription initiation assays were performed in transcription buffer (40mM Hepes pH 7.3, 250mM NaCl, 2.5mM MgCl$_2$, 0.1mM EDTA, 0.1mM ZnSO$_4$), 0.1µg/µl BSA and 1mM DTT using 10nM gdh-C20 template DNA, 10nM RNAP, 238nM TBP and 135nM TFB in a total volume of 25µl. Samples were incubated with 40µM GpU (Guanylyl-5´-phosphatidyl-Uracil) dinucleotide primer and 74kBq [α^{32}P]-UTP (111TBq/mmol) for 10 minutes at 70°C. RNAs were isolated with 40µl PCI and vigorous mixing for 40 seconds. The samples were centrifuged for 5 minutes at 21.000g and supernatant was diluted in 3x loading dye (98% v/v formamide, 40µM EDTA, 0.025% w/v bromophenol blue and 0.025% w/v xylene cyanol) and separated on a 7M urea-28% acrylamide gel.

3.   Run-off transcription assay

Transcription assays were performed using 1xTB-0 transcription buffer (40mM Hepes pH 7.3, 250mM NaCl, 2.5mM MgCl$_2$, 0.1mM EDTA, 0.1mM ZnSO$_4$, 0.1µg/µl BSA and 1mM DTT). 10nM gdh-C20 template DNA, 10nM RNAP, 238nM TBP and 135nM TFB were incubated with 440µM GTP, 440µM ATP and 440µM CTP together with 2.7µM UTP and 49kBq [α^{32}P]-UTP (111TBq/mmol) for 10 minutes at 70°C in a total volume of 25µl. RNAs were isolated with 40µl PCI and vigorous mixing for 40 seconds. The samples were centrifuged at 21.000g for 5 minutes and supernatant was diluted in 3x loading dye (98% formamide, 40µM EDTA, 0.025% w/v bromophenol blue and 0.025% w/v xylene cyanol) and separated on a 7M urea-8% acrylamide gel.

4.   Chase experiments and stalled transcription complexes

Transcription assays were performed as described in (III.C.3) but a nucleotide mix containing 40µM ATP, 40µM GTP, 2.5µM UTP and 49kBq [α^{32}P]-UTP (111TBq/mmol) were used. Samples were incubated at 80°C for 3 minutes, whereas the reactions were stopped by cooling down the cycler, or chased by adding a NTP-mix containing 440µM of each NTP without radiolabeled UTP and further incubated at 80°C for additional 10 minutes. RNAs were isolated with 40µl PCI and vigorous mixing for 40 seconds. The samples were centrifuged at 21.000g for 5 minutes and supernatant was diluted in 3x loading dye, whereas short transcripts with a length of 6-25nt were separated on a 7M urea 28% polyacrylamide gel, for longer transcripts a 7M urea-8% acrylamide gel was used.

5.   Potassium permanganate footprinting

Reactions were performed in 0.5xTB-0 instead of 1xTB-0 to reduce Na-HEPES concentration, as this buffer reagent increased the reactive potential which caused high background on the gel (Lozinski, Wierzchowski 2003). Footprints were performed with reduced buffer conditions, 0.1µg/µl BSA, 1mM DTT and 3µl of labeled DNA (III. A. 2), 10nM RNAP, 238nM TBP and 135nM TFB and 217nM TFE in a total volume of 25µl. Samples were incubated at 70°C for 10 minutes, then 2.5µl of a 250mM KMnO$_4$ solution was added and further incubated for 5 minutes. Reactions were stopped using 1.5µl β-mercaptoethanol and 20µl stop solution (100mM EDTA, 1% w/v SDS) and tubes were subsequently placed on ice. Samples were placed on the MPS and the supernatant was discarded. Beads were resolved in 18µl H$_2$O and mixed with 2µl piperidine. After incubation at 90°C for 30 minutes

the supernatant was transferred into 40µl PCI and mixed for 40 seconds. After centrifugation at 21.000g for 5 minutes 17µl of the upper phase was transferred into a fresh tube, then 30µl H$_2$O, 5µl 3M NaAc pH 5.3, 1µl glycogen (20mg/ml) and 125µl ethanol (>99.98%) was added. The samples were mixed and incubated at -80°C for 45 minutes, and centrifuged at 21.000g for 30 minutes. The supernatant was removed and the pellet was washed with 800µl ethanol (70%), and centrifuged for 10 minutes at 21.000g. The pellet was dehumidified using a vacuum concentrator (Concentrator 5301, Eppendorf). DNA was dissolved in 10µl TE´-buffer and mixed with 5µl 3x loading dye and incubated at 95°C for 3 minutes. Samples were separated on a 6% sequencing gel.

6. Crosslinking experiments

The method was modified form (Micorescu et al. 2008). If not otherwise noted, UV-induced photo crosslink reactions were performed in a total volume of 25µl with 2nM of radiolabeled DNA under consideration of the half time of the [α^{32}P] dNTP, 13nM RNAP, 238nM TBP and 135nM TFB (or its variants) in transcription buffer (40mM Na-HEPES pH 7.3, 250mM NaCl, 2.5mM MgCl$_2$, 0.1mM EDTA, 0.1mM ZnSO$_4$) with 0.1µg/µl BSA and 1mM DTT and, depending on the reaction, with additional 217nM TFE and/or NTP Mix (40µM ATP, 40µM GTP and 2.7µM UTP). The samples were covered with mineral oil to prevent evaporation and incubated 5 minutes at 80°C. 2µg heparin was added as competitor and the samples were exposed to UV-light (Philips TUV 15W/G15T8 UV-C) for 20 minutes at 80°C. Despite the recommended wavelength of 350-360nm (Kauer et al. 1986; Dorman, Prestwich 1994), crosslinking efficiency was determined to work optimally at a wavelength of λ=300nm in our system. 15µl of the samples were recovered from oil and digested with 1.65µl DNase I-Mix (1U DNase I, 40mM TRIS/HCl pH 8.0, 7mM MgCl$_2$, 100mM NaCl, 5mM CaCl, 1mM DTT, 0.1µg/µl BSA, 1mM PMSF, 1µg/ml pepstatin, 1µg/ml leupeptin, 50% (v/v) glycerol) for 10 minutes at 37°C. 1.0µl of 10% (w/v) SDS solution was added and incubated at 90°C for 3 minutes. Samples were mixed with 1.92µl 12x Zn/HoAc-Mix (0.1mM ZnSO$_4$, 12% (v/v) glacial acetic acid) and 1.4µl S1 nuclease-mix (10U S1 nuclease, 20mM TRIS/HCl pH 7.5, 50mM NaCl, 0.1mM ZnSO$_4$, 0.1µg/µl BSA, 1mM PMSF, 1µg/ml pepstatin, 1µg/ml leupeptin and 50% (v/v) glycerol) and incubated for 10 minutes at 37°C. Reactions were denatured with 4.11µl 6x SDS loading dye, denatured and completely separated using 12% SDS-PAGE. The gel was transferred into fixation solution (30% (v/v) ethanol, 10% (v/v) glacial acetic acid) and incubated over night to reduce background signals. The gel was fixated to whatman paper and radioactive signals were detected with Phospho Imager (FUJIFILM FLA5000) and analyzed with AIDA v4.27 software.

D. FRET measurements and data acquisition

FRET (Förster resonance energy transfer) measurements were performed by Kevin Kramm using confocal microscopy and TIRF microscopy (Total internal reflection fluorescence). For these experiments a SSV T6 promoter DNA containing a 5´ biotin on the nt-strand, an ATTO647n acceptor dye on the 5´ t-strand, and an internal Cy3b donor dye next to the B responsive element was used (Gietl et al. 2014). Samples were prepared with 20pM SSV T6 promoter DNA and 1µM TBP and 1µM TFB in transcription buffer containing 40mM Tris/Cl pH 7.5, 250mM NaCl, 0.1mM EDTA pH 8.0, 2.5mM MgCl$_2$ and 0.1mM ZnSO$_4$ in a total volume of 200µl. Measurements were performed with the respective microscope technique at room temperature by Kevin Kramm. Three datasets of the confocal microscopy were obtained as absolute values for FRET efficiency (E) and the corresponding absolute values for the number of counted events. The efficiency E can be defined as the number of energy transfer events during a donor excitation event. The values of each dataset were binned to

50 data points in total and were plotted as histograms and fitted with a Gaussian fit, including the calculated mean and standard deviation. Raw data of the TIRF microscopy measurements were analyzed with iSMS software. In total 250 traces of co-localized FRET pairs of each sample (DNA without factors and DNA with TFB and TBP) were analyzed and bleaching events of the donor dye were eliminated. Each dataset was binned with a width of 1.18 and plotted as a histogram. The datasets were fitted with a Gaussian fit, and standard deviation as well as the mean FRET efficiency were calculated by statistical operations with SigmaPlot software and/or alternatively with MATLAB (integrated in iSMS software package).

IV. Results

The first part of this chapter is about RPA, in which the results of the *in vitro* experiments are shown. Results concerning TFB are structured in three sections. The first is about FRET-measurements and DNA bending, and the second part is about the role of the TFB B-reader loop and its charge-dependent interactions during transcription initiation. The last section is the main part of this thesis, and the topology of TFB in the preinitiation complex as well as general structural rearrangements of TFB during transcription initiation was investigated. All sections are discussed in a separate chapter (V. Discussion).

## A. Analysis of the replication protein A during transcription

Thomas Fouqueau, a former PhD student at our institute, showed that *Pab*RPA stimulates transcription *in vitro* (Pluchon et al. 2013). However, details about the mechanism of the observed stimulation lacked for the replication protein A. To gain more insights into molecular interactions between the related *Pfu*RPA and the transcription machinery of *P. furiosus,* the protein was analyzed using different *in vitro* assays. The cell cultivation of the respective genetic modified *P. furiosus* strain, as well as purification of *Pfu*RPA was performed by Julia Winter during an internship at our institute in 2013. Using SDS-PAGE, three subunits were observed at the expected height of 14kDa, 32kDa and 41kDa and additional MS data confirmed the presence of the heterotrimeric protein (data not shown).

1. RPA in transcription initiation

To test the functionality of the protein, an electro mobility shift assay (EMSA) was performed to show the predicted preference to single stranded DNA (Figure 9 A), which is a typical feature of SSB proteins. *Pfu*RPA was added to the samples with increasing concentrations and a shift was only observed in presence of single-stranded DNA (Figure 9 A, lane 2-4), indicating interaction of RPA with the template. Here, the signal intensity rose with increasing amounts of RPA to 130% at 348nM, and 140% at 523nM, respectively, in comparison to 174nM RPA. In contrast, samples with dsDNA did not show a specific shift on the gel except some unspecific signals. The dsDNA template was generated by hybridization of the 5´-Cy3-labeled ssDNA with the complementary unlabeled strand. Unspecific signals were located on the same height as the signals of samples where ssDNA was present, and the signal intensity of the lanes with increasing amounts of RPA did not differ, indicating that RPA was bound to a small subpopulation of residual non-hybridized ssDNA templates. In the lanes where RPA was absent, no shift was observed for ssDNA, but for dsDNA, which might be a result of a deficient loading of the sample containing ssDNA at 523nM RPA. However, the results demonstrated the expected preference to single-stranded DNA.

To analyze if RPA is part of the preinitiation complex, several methods like EMSA with fluorescently and radioactively labeled DNA as well as the more sensitive western blot approach was used (data not shown). As the results of these experiments did not indicate a presence of RPA in the preinitiation complex, a more functional approach was used to identify a role of RPA in transcription initiation. The results of the abortive transcription assays are summarized in figure 9 B. To analyze the impact of RPA on the first phosphodiester bond formation of an initiating complex, three independent technical replicates were analyzed with equalized buffer conditions and increasing RPA concentrations. The signal intensity derived from the radiolabeled 3nt RNA products did not differ with increasing concentrations of RPA in comparison to the sample without RPA, indicating that RPA does not influence the formation of the first phosphodiester bond. Therefore, RPA does not stimulate RNA synthesis in the initiation stage of transcription.
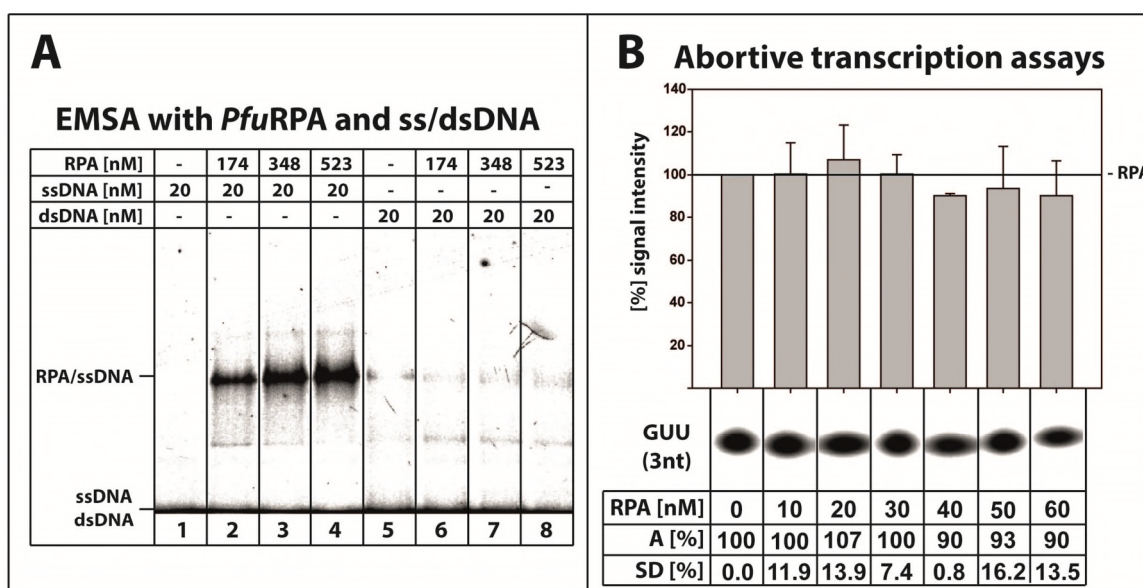
Figure 9: Analysis of RPA during transcription initiation and its preference to ssDNA. **A)** EMSA with different concentrations of RPA with ssDNA and dsDNA confirmed the specific interaction to ssDNA. **B)** Abortive transcription assays with increasing concentrations of RPA did not show effects, indicating that RPA functions not during the first phosphodiester bond formation.

2.   RPA in transcription elongation

To confirm the results of (Pluchon et al. 2013), an *in vitro* run-off transcription assay was performed with increasing concentrations of the replication protein A (Figure 10 A). Under consideration of the buffer conditions, the signal intensity of the formed 113nt run-off transcripts rose with increasing amounts of RPA, indicating that more radiolabeled transcripts were formed in presence of RPA. The activation of transcription was calculated to 2.9-fold in average of three independently performed experiments with 50nM RPA used. This was also observed for the *P. abyssi* RPA with the same amounts, showing that the *P. furiosus* RPA has the same effect under run-off conditions. These results indicated that RPA functions during elongation of transcription, as experiments with respect to the initiation did not show an effect in presence of RPA. To investigate the possible role in transcription elongation, a 4kb plasmid containing a glutamate dehydrogenase (gdh) promoter and a stalling site at position +45 relative to the TSS was used. Complexes were stalled at this register using NTP mix containing [α-^{32}P] UTP, but no CTP (Figure 10 B, lane 1, 3). RPA-buffer (lane 3) and 10nM RPA (lane 4) were added to reactions for one additional minute. The complexes where then chased for 10 minutes by addition of a molar excess of unlabeled UTP and CTP, whereby the radiolabeled 45nt RNAs were extended. The results showed that in absence of RPA the ~4kb run-off product was formed, but several intermediates (marked with *) appeared. These intermediates derived from chased complexes and indicating an accumulation of these transcripts during elongation of transcription possibly due to a drop-off of the RNA polymerase from the plasmid or internal pausing sites. In contrast, the presence of only 10nM RPA could prevent the formation of these intermediates. In addition, the signal intensity of the ~4kb band was also higher than without RPA, indicating that more complete transcripts were formed under these conditions. This result leads to the assumption that RPA functions during elongation possibly by stabilizing the RNA polymerase during RNA synthesis. This interaction might reduce stalling and pausing events of the polymerase which occur regularly during transcription, and therefore more transcripts were formed within the same time in comparison to samples without RPA. In run-off transcription assays, a

significant change of the pattern of shortened intermediates between samples where RPA is absent or present was not observed (data not shown). Therefore, another suggestion is that RPA increase the transcription speed of the RNA polymerase, which would also lead to a higher number of formed transcripts within the same time frame.
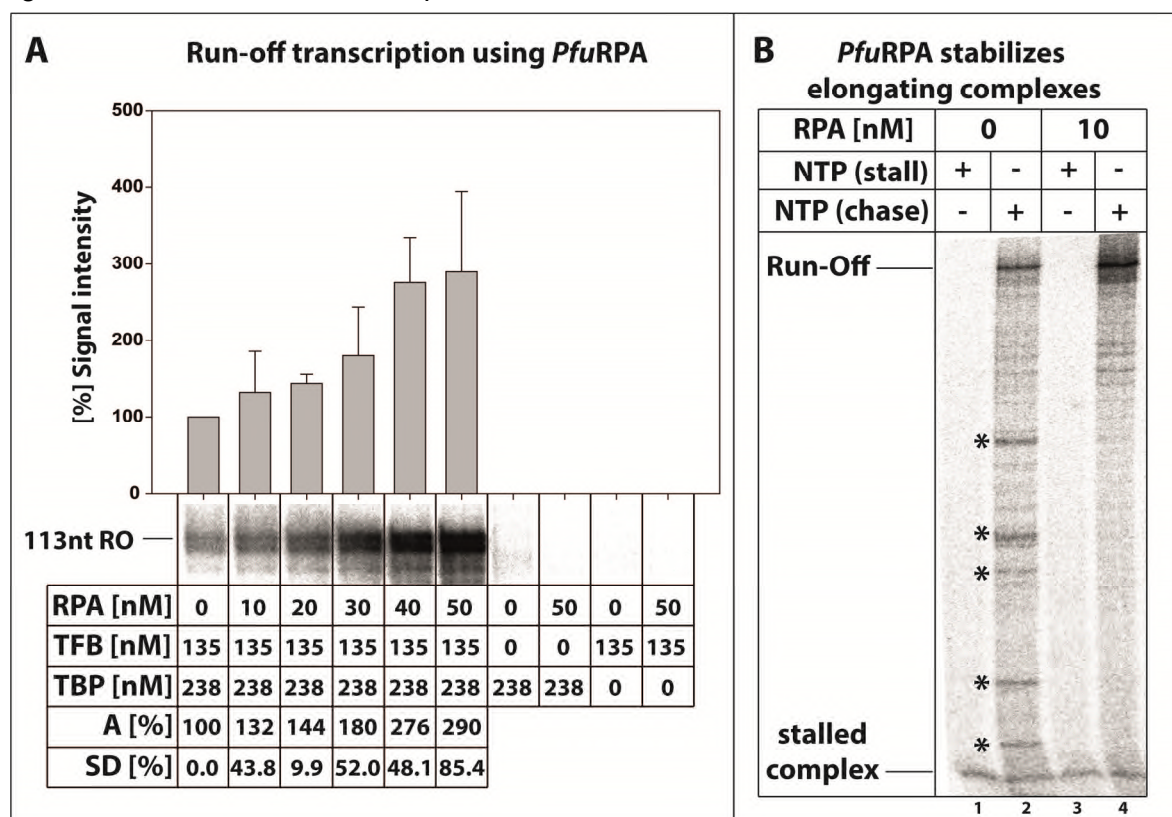
Figure 10: RPA functions during elongation of transcription. **A)** Run-off transcription assay with increasing concentrations of RPA revealed the stimulatory effect, as the formation of the 113nt run-off transcript increased with higher concentrations of RPA. **B)** Chase experiments of stalled elongation complexes on plasmids showed intermediates (marked with *) in absence of RPA, whereas in presence of RPA no signals at the respective heights were observed, indicating a stabilization effect of RPA during elongation.

To investigate differences in the transcription speed, which is the number of incorporated nucleotides per time, transcription assays were performed with and without RPA and reactions were stopped after distinct time points (Figure 11). For this experiment the run-off length of the regular gdh-C45 template was extended from 113nt to 250nt by creating new templates. Complexes were stalled at register +45 in absence and presence of RPA on the linearized template, and chased by addition of a nucleotide mix containing unlabeled UTP and CTP in a molar excess. In a first assay, reactions were stopped after every 5 seconds (Figure 11). Full run-off transcripts were observed after 20 seconds in absence of RPA. In contrast, the polymerase need only 15 seconds for the 250nt transcript length in presence of RPA. Comparison of the shorter intermediates (Figure 11, red asterisks) also showed that in presence of RPA formation of transcripts occurred much faster than in absence of RPA. The results demonstrated that the stalled RNA polymerase can transcribe much faster in presence of RPA. However, the same time-dependent experiments were performed without stalled complexes, in which preinitiation complexes were incubated with RPA and NTPs and reactions were stopped after distinct time steps (data not shown). In contrast, the differences between formed transcripts in absence or presence of RPA did not differ remarkably, indicating that the transcription speed of the RNAP is only slightly increased in presence of

RPA. This finding leads to the assumption, that RPA can help the RNAP to restart transcription from stalled complexes, and slightly increases the processivity of the RNA polymerase.
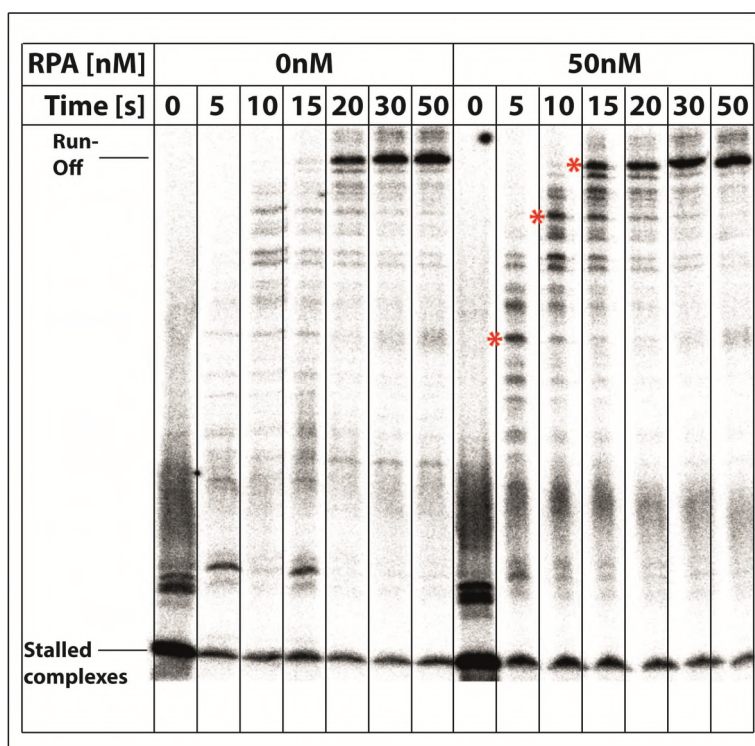
Figure 11: Time-dependent transcription reactions of chased complexes in absence and presence of RPA. Stalled complexes were chased and stopped after distinct seconds. In presence of RPA the 250nt run-off product was formed at 15 seconds, whereas in absence of RPA the run-off product was formed after 20 seconds.

3.   Summary of *Pfu*RPA experiments

The purified heterotrimeric *Pfu*RPA showed the expected preference to single stranded DNA as it was expected for members of the SSB protein family (Figure 9 A). *In vitro* run-off transcription assays further confirmed the observed effect of *Pab*RPA, in which the formation of transcripts is stimulated up to 2.9-fold (Pluchon et al. 2013), indicating that *Pfu*RPA acts in a similar manner like RPA of *P. abyssi* (Figure 10 A). However, the exact mechanism of this stimulation is unknown, and therefore experiments were performed to unravel its possible function. Western blot experiments as well as EMSAs with radio labeled and fluorescently labeled DNA templates indicated that RPA is not part of the initiation complex (data not shown). This finding is further supported by the fact that first phosphodiester bond formation assays (Figure 9 B) did not show effects in presence of RPA in comparison to samples without RPA. Moreover, RPA functions in the elongation stage of transcription. It could be demonstrated that elongating RNA polymerases were stabilized in transcription reactions using a complete plasmid (Figure 10 B), and formation of intermediates, which appeared in absence of RPA, can be prevented in presence of this protein. However, it could not be determined if these intermediates are internal pausing sites, or the polymerase loses its interaction to the DNA template. In both cases RPA would help the polymerase to overcome the pausing sites or increase the stability of the elongating complex to prevent its dissociation from the template. Chase experiments with stalled complexes further showed that in presence of RPA the polymerase is able to restart transcription much faster than in absence of RPA (Figure 11).

## B. DNA bending experiments of *P. furiosus* TFB using FRET

The experiments with RPA focused on the elongation of transcription. However, this chapter is about the first steps in transcription initiation, in which TBP and TFB associate to the promoter of the DNA. The resulting DNA bending is a prerequisite for the formation of a preinitiation complex (Nikolov et al. 1995). In the euryarchaeon *M. jannaschii* it was shown that TBP alone can bend DNA, whereas in *S. acidocaldarius*, a member of the Crenarchaeota, DNA bending requires TFB (Gietl et al. 2014). To investigate DNA bending in the organism *P. furiosus*, single molecule analysis was performed using confocal and total internal reflection fluorescence (TIRF) microscopy techniques with kindly support of Kevin Kramm, PhD student of the laboratory of Prof. Dina Grohmann. To enable comparison with results described in literature, the Sulfolobus spindle-shaped virus 1 T6 (SSVT6) promoter was used as described in (Gietl et al. 2014), together with own *Pfu*TBP and *Pfu*TFB. This short template contains a TATA box and a BRE to enable binding of TBP and TFB to the DNA, and a biotin is present at the 5´ end of the nt-strand for immobilization of the DNA on a surface (Figure 12 C).
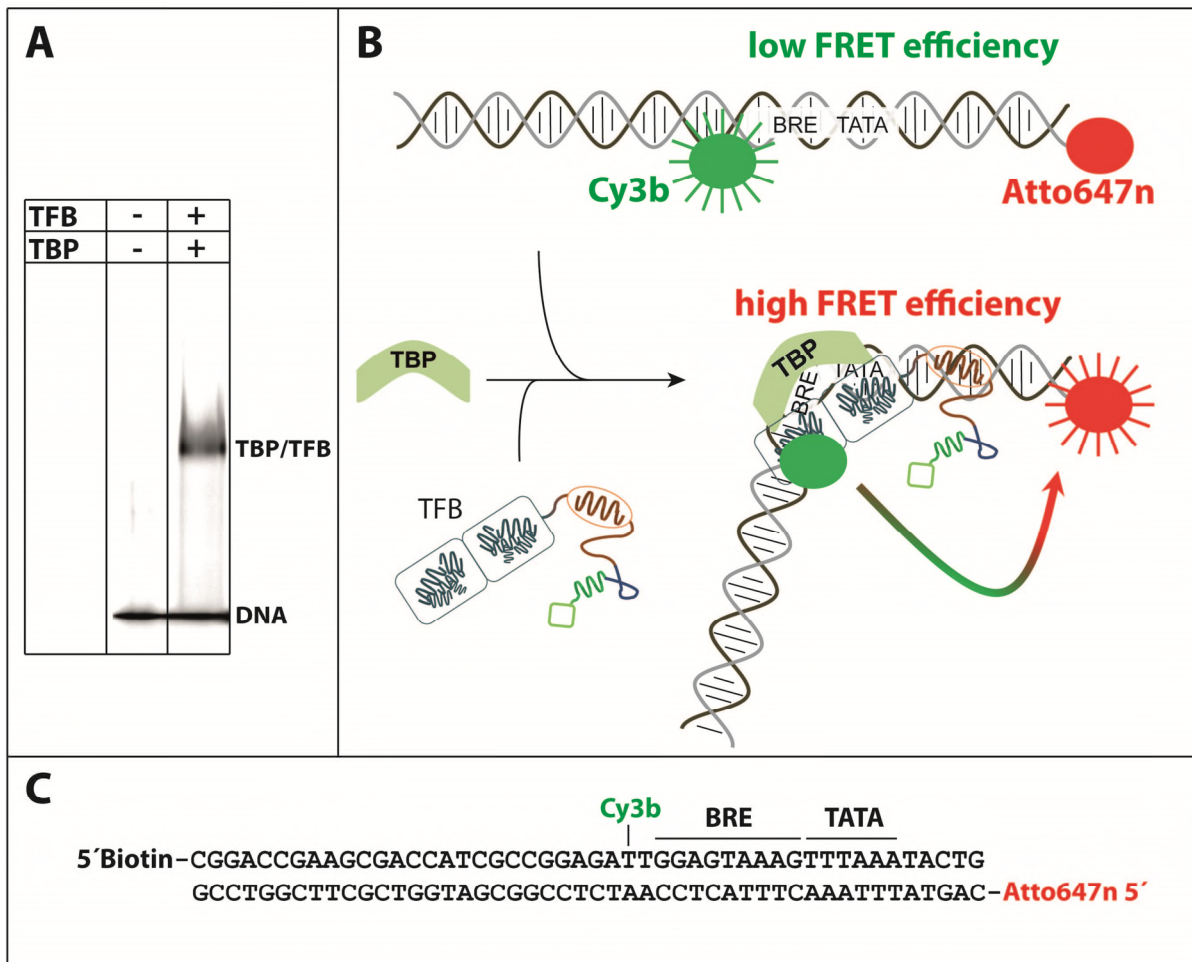
Figure 12: EMSA, principle of FRET and SSVT6 template overview. **A)** Specific shift signal was observed in presence of *Pyrococcus* TBP and TFB, demonstrating that the proteins bind to the viral promoter. **B)** Principle of the FRET measurement. In the unbent state of DNA the distance between the donor dye Cy3b and the acceptor molecule Atto647n is higher than in the bent state of DNA. In the bent state, the FRET efficiency is higher than in the unbent state due to a higher number of energy transfers caused by a shorter distance. **C)** Representation of the SSVT6 promoter used in the study. The template contains a 5´ biotin for immobilization, a Cy3b donor dye next to the BRE, and an acceptor molecule Atto647n at the 5´ t-strand (Gietl et al. 2014).

In addition, two fluorescent dyes are fused to the template, one donor dye Cy3b (Excitation Abs-λ_{max}= 558nm; Emission Em-λ_{max}=572nm) located next to the BRE on the nt-strand, and an acceptor molecule Atto647n (Excitation: Abs-λ_{max}= 644nm; Emission: Em-λ_{max}=669nm) on the 5´ end of the t-strand. The donor dye (D) is excited by a laser beam of a distinct wavelength under the microscope, whereas the emission spectrum of this donor overlaps with the excitation spectrum of the acceptor (A) molecule. If both D and A are in distance closer than 10nm, the acceptor emits light of a distinct wavelength if excited by the donor dye due to Förster resonance energy transfer (FRET) (Figure 12 B). The number of energy transfers over the number of donor excitations is termed the FRET efficiency (E). A higher E value indicate a shorter distance between D and A. Therefore, if DNA is bent, the FRET efficiency increases due to the shorter distance between the donor and acceptor dye.

First, the binding of TBP and TFB of *P. furiosus* to the SSVT6 promoter was tested in an EMSA (Figure 12 A). In presence of TBP and TFB a specific band appeared on the gel, indicating the formation of a ternary complex consisting of DNA and the two transcription factors. Because the usual buffer conditions were insufficient for a proper measurement in the fluorescence microscopy, the composition of the buffer was changed. Reducing agents like DTT were eliminated and the interfering buffer substance Na-Hepes was exchanged with Tris. *In vitro* transcription reactions showed no difference between both buffer conditions (data not shown). First results were obtained from the confocal microscopy. The great difference between confocal and TIRF microscopy is the immobilization of the DNA to a polyethylene glycol surface in the TIRF (Gietl et al. 2014), where single molecules can be visualized as small spots. In contrast, samples in the confocal microscopy just diffuse through a focused area. In total, three measurements were performed: One contained only the SSVT6 promoter, in a second sample TBP and the template were present, and in a third sample additional TFB was added to the template and TBP (Figure 13). The sample with DNA (Figure 13 A) showed a stable conformation with a mean FRET efficiency of 28.4% ± 10.6%. This value represented the unbent state of DNA, as no other proteins were present in the reaction. In the sample containing additional TBP, the mean FRET efficiency did not differ and had a mean of 28.0% ± 10.5% (Figure 13 B).
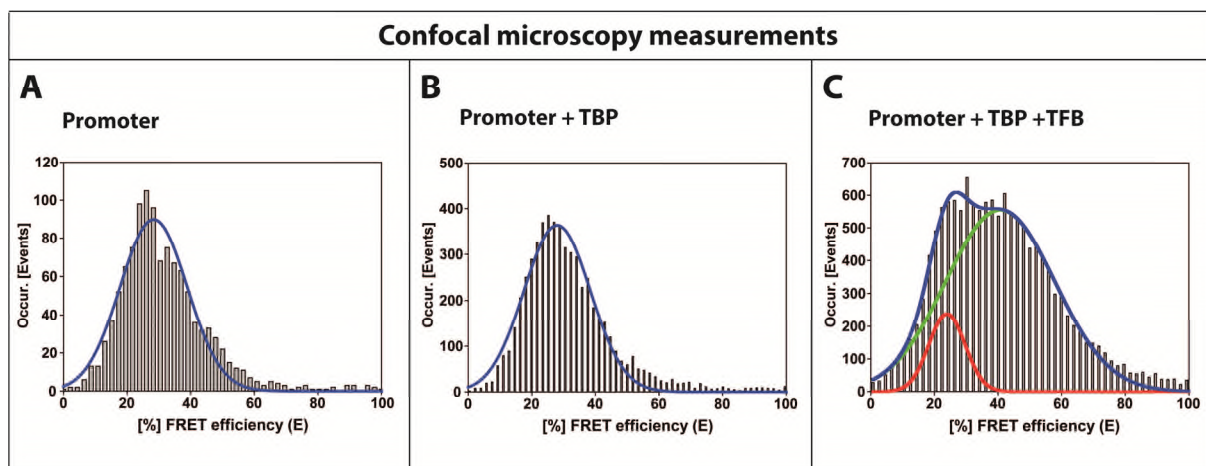
Figure 13: Results of the confocal microscopy measurements. **A)** Measurement of SSV T6 promoter only. DNA shows a stable conformation and a mean FRET efficiency of 28.4% ± 10.6%. **B)** Addition of TBP to DNA does not change the mean FRET efficiency (28.0% ± 10.5%), indicating no bending effect for TBP alone. **C)** TFB induces DNA bending. Two populations are present (blue double Gaussian curve): a first population with a mean FRET efficiency of 24.0% ± 5.7% (red curve) representing the unbent conformation, and a high FRET population with a mean of 40.7% ± 17.2%, indicating DNA bending (green curve).

This finding suggests that TBP alone is not able to bend DNA, or keep the DNA in a bent state. Addition of TFB leads to a second population with an increased mean FRET efficiency of 40.7% ± 17.2% (Figure 13 C). This indicates a bent state for the DNA, as the distance between donor and acceptor is decreased. A second population was detected with a mean FRET efficiency of 24.0% ± 5.7%, indicating that a subpopulation of DNA in an unbent conformation is still present. The results of the confocal microscope show that DNA bending in *P. furiosus* depends on the presence of TFB, whereas either TBP and TFB bind and bend DNA simultaneously, or binding of TFB stabilizes the TBP-DNA interaction and bend DNA.
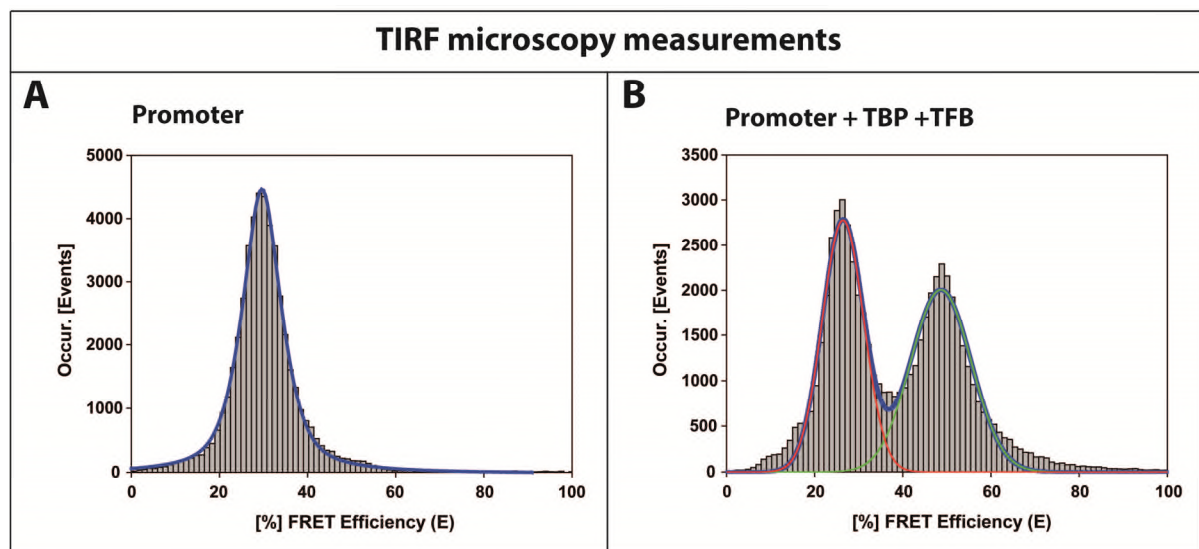

Figure 14: Results of the TIRF microscopy measurements. **A)** The same setup was used as for confocal microscopy measurements, but DNA and the respective complexes were immobilized on a surface. DNA showed a stable conformation with a mean FRET efficiency of 29.6% ± 0.7%. **B)** Measurement with additional TFB and TBP. The overall distribution of the FRET efficiency showed two maxima (blue Gaussian fit), one with a mean FRET efficiency of 26.4% ± 0.2% (red Gaussian fit) indicating DNA in the unbent state, and a second with a mean FRET efficiency of 48.8% ± 0.2% (green Gaussian fit) indicating a DNA bending in presence of TFB and TBP.

To confirm these results with a second method, TIRF microscopy measurements were performed in absence and presence of both transcription factors. To determine the mean FRET efficiency for the unbent state, a sample containing only DNA was immobilized on the surface and measured. Here, FRET efficiency had a mean value of 29.6% ± 0.7% (Figure 14 A). Addition of TFB and TBP to the template leads to a mixture of two populations, one unbent conformation with a mean E = 26.4% ± 0.2%, whereas the second mean E = 48.8% ± 0.2% (Figure 14 B). This result confirmed the measurements of the confocal microscopy and further indicates that TFB induces bending of DNA, whereas TBP is not able to bend DNA alone.

C. The role of the TFB B-reader loop in transcription initiation

At the beginning of this PhD thesis, a crystal structure of the eukaryotic yeast initially transcribing complex was published containing *Sc*TFIIB, RNAP II, DNA and a 6nt RNA bound to the transcribed strand at a resolution of 3.4Å (Sainsbury et al. 2013). The structure revealed that the tip of the B-reader loop of *Sc*TFIIB is located in close proximity to the active site of the polymerase and might interact with the nascent RNA at a length of 6 nucleotides (Figure 15 A). Based on this finding, it was hypothesized that the advancing RNA is separated from DNA with the support of the *Sc*TFIIB B-reader loop domain. Due to the fact, that the respective site of *Sc*TFIIB contains two aspartate residues D74 and D75 (*Pfu*TFB E53 and R54), it was suggested that separation occurs via charge-dependent interactions between the negative charged amino acids of *Sc*TFIIB B-reader loop and the 5´end of the nascent RNA chain. The separation of RNA from the template is necessary for transcription initiation, because RNA has to be guided towards the exit channel of the RNA polymerase to enter productive elongation (Figure 15 C).
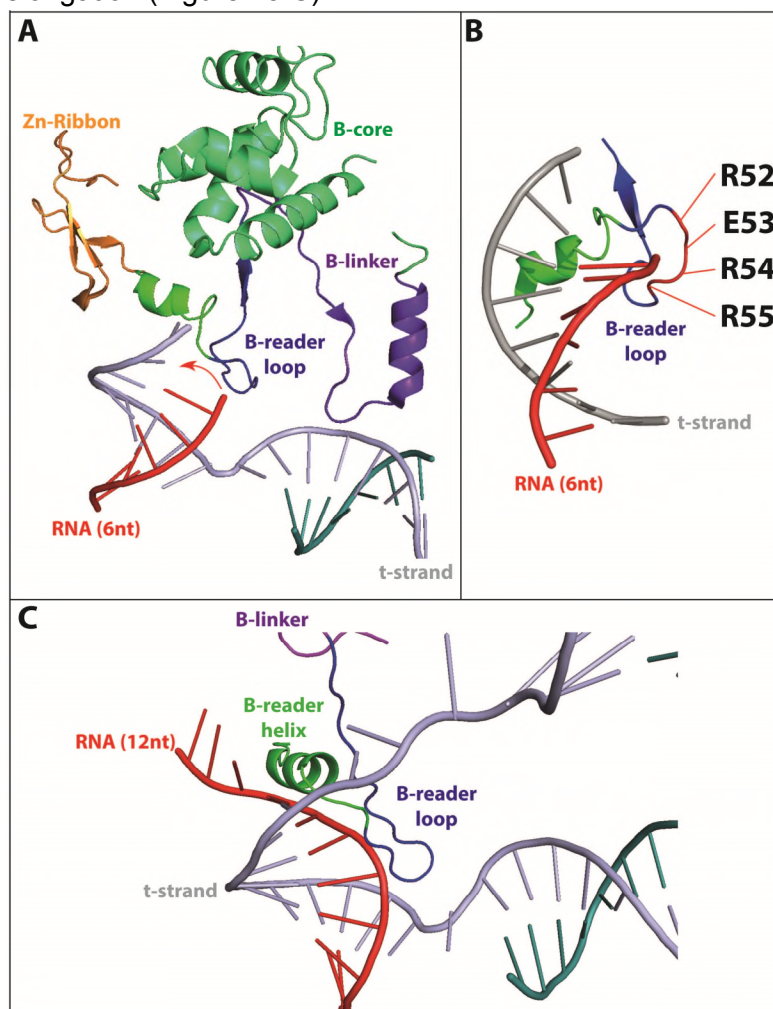
Figure 15: Separation model based on the published crystal structure 4BBS. Modified from (Sainsbury et al. 2013). **A)** The nascent RNA (red) interacts with the B-reader loop domain of TFB (blue) to be separated from the transcribed strand (grey) and guided towards the exit channel. Other domains of TFB are shown in green (B-reader helix), dark green (B-core), orange (Zn-ribbon) and purple (B-linker). **B)** Locations of selected amino acids for alanine substitution are indicated and directly interact with nascent RNA in the structure (4BBS). **C)** Model of the separated RNA. A 12mer RNA was modelled into the structure based on the predictions by (Sainsbury et al. 2013). RNA is separated by the B-reader loop and is guided towards the exit channel, whereas it is closely located to the B-reader helix. Color code is the same as in A and B.

Comparison of the eukaryotic B-reader loop tip of different organisms showed that the respective site contains usually one negative charged amino acid in all selected organisms except *M. jannaschii*, whereas archaea additionally possess an overall positive charge due to the presence of basic amino acids (Figure 16). Eukaryotic organisms lack positive charged amino acids at the corresponding site. To elucidate the specific role of the acidic amino acid and in general the B-reader loop of *P. furiosus* with respect to RNA-DNA separation and its role in archaeal transcription, several *Pfu*TFB alanine variants were created. The charge of the loop was stepwise eliminated (Figure 15 B). In total, four single mutations R52A - R55A, three double variants R52E53A, E53R54A, and R54R55A, and a complete loop alanine substitution (R52-R55A; referred to as LoopA) were developed during the Bachelor thesis of C. Dick under my attendance in 2014. In contrast to the *Sc*TFIIB B-reader loop, the overall charge of the *P. furiosus* B-reader loop is positive (Figure 16). Nevertheless, its role in transcription was analyzed using different *in vitro* transcription assays.

Figure 16: Charge distribution of the conserved TFIIB/TFB B-reader domain of different organisms. Sequences of members of the eukaryotic domain are compared with sequences derived from archaeal organisms (NA=Nanoarchaeota; CA=Crenarchaeota; EA=Euryarchaeota). The B-reader loop tip is indicated by the red line, whereas the location of the B-reader helix (green line) and the loop (blue line) is shown. Amino acids in grey boxes are unpolar, in white boxes neutral, in blue boxes positively charged, and in red boxes acidic. The multiple sequence alignment was performed using ClustalOmega.

1.   Analysis of TFB Alanine substitutions in transcription assays

The ability of the developed TFB variants to form a preinitiation complex was tested in an EMSA experiment (Figure 17). Signals for the ternary TFB/TBP/DNA-complex occurred only in presence of both transcription factors (Figure 17 A, lane 4), whereas single factors present with DNA do not form an unspecific interaction (Figure 17 A, lane 2, 3). Addition of RNAP leads to signals located higher on the gel (Figure 17 A, lane 5-9; B, lane 3-7), representing preinitiation complexes. Two separated bands were detected in samples containing RNAP, possibly as a result of two PIC populations. It is possible that one population contains the stalk subunits Rpo4/7, and the other population lacks the stalk. The experiment showed that preinitiation complexes formed with single alanine substitutions (Figure 17 A) have the same pattern like the wild type TFB (wtTFB), except TFB-E53A. The ternary TFB-E53A/TBP/DNA-complex runs lower than the other complexes on the gel. Therefore the electromobilic property differs in comparison to usual ternary complexes. The double alanine substitutions

and the loop tip substitution showed no altered patterns in comparison to the wtTFB (Figure 17 B), indicating that all TFB variants are able to form preinitiation complexes.
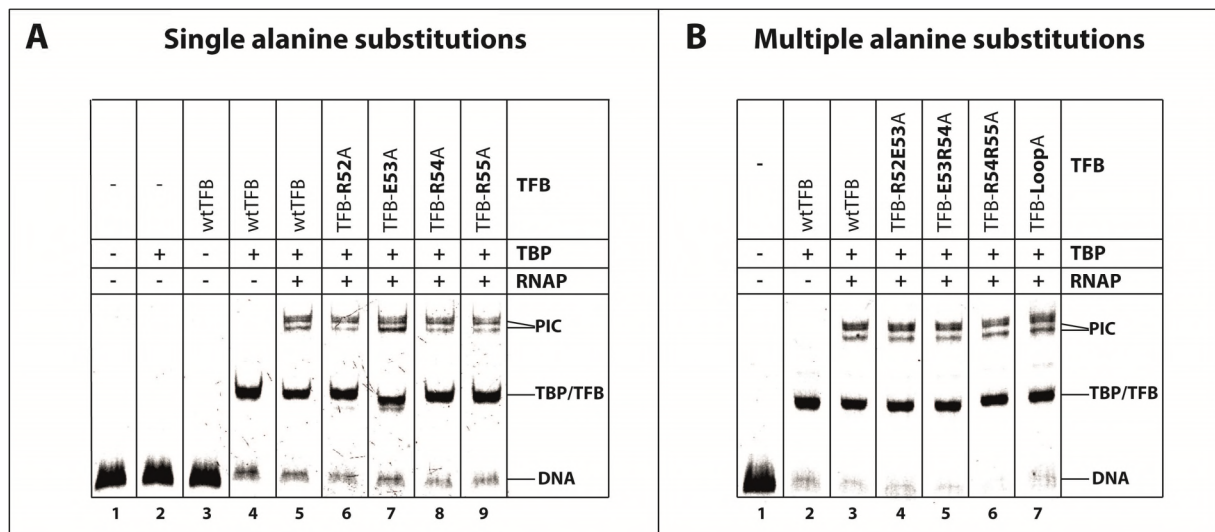
Figure 17: EMSA on a 5% native gel of TFB alanine variants and their ability to form a preinitiation complex. **A)** A specific shift was observed only in presence of both transcription factors or with additional RNAP (lane 4-5). TFB variants containing single alanine substitutions form preinitiation complexes (lane 6-9) in the same manner like the wild type TFB (lane 5), whereas TFB-E53A showed an altered pattern as the TFB/TBP/DNA-complex runs lower than the comparable complexes. Two bands were detected in presence of RNAP, possibly representing two PIC populations with and without the stalk subunits Rpo4/7. **B)** TFB variants containing multiple alanine substitutions showed the same pattern like the wtTFB, indicating the formation of a preinitiation complex.

To test the ability of the formation of the first phosphodiester bond of the RNA polymerase initiated with the TFB variants, an abortive transcription assay was performed (Figure 18 A). In this assay transcription is initiated using a primer consisting of two nucleotides GpU, and $[\alpha^{32}P]$ UTP, whereas the primer is extended to a 3nt radiolabeled product by the RNA polymerase to measure the capability of the RNAP to form the first phosphodiester bond. The results of the experiments showed that TFB-R52A and TFB-E53R54A showed moderate transcription levels in comparison to the wtTFB, indicating that the elimination of these charges at the respective sites do not inhibit or stimulate transcription initiation. In contrast, RNAPs initiated with TFB-R54A showed a markedly reduced number of aborted transcripts of only 32% in comparison to wtTFB, suggesting insufficient capability of the enzyme to form a phosphodiester bond. Single alanine substitutions TFB-E53A and TFB-R55A, the double alanine variant TFB-R54R55A and the TFB-LoopA mutation showed almost no transcription initiation, indicating that exchange of these amino acids with alanine are not able to initiate transcription correctly. Surprisingly, the double substitution R52E53A showed a two-fold increase of the signal intensity in comparison to the wtTFB. This finding suggests that the two alanine residues at these positions stimulate the formation of the first phosphodiester bond. To investigate the impact of the TFB variants on the formation of a run-off transcript, a multiple-round *in vitro* transcription assay was performed (Figure 18 B). All transcription relevant components were incubated with all four nucleotides and $[\alpha^{32}P]$ UTP, and radiolabeled transcripts with a length of 113nt were formed. The single substitutions TFB-R52A, TFB-R54A and TFB-E53R54A showed almost similar results like in the abortive transcription assay, whereas the number of transcripts is slightly reduced to 50%
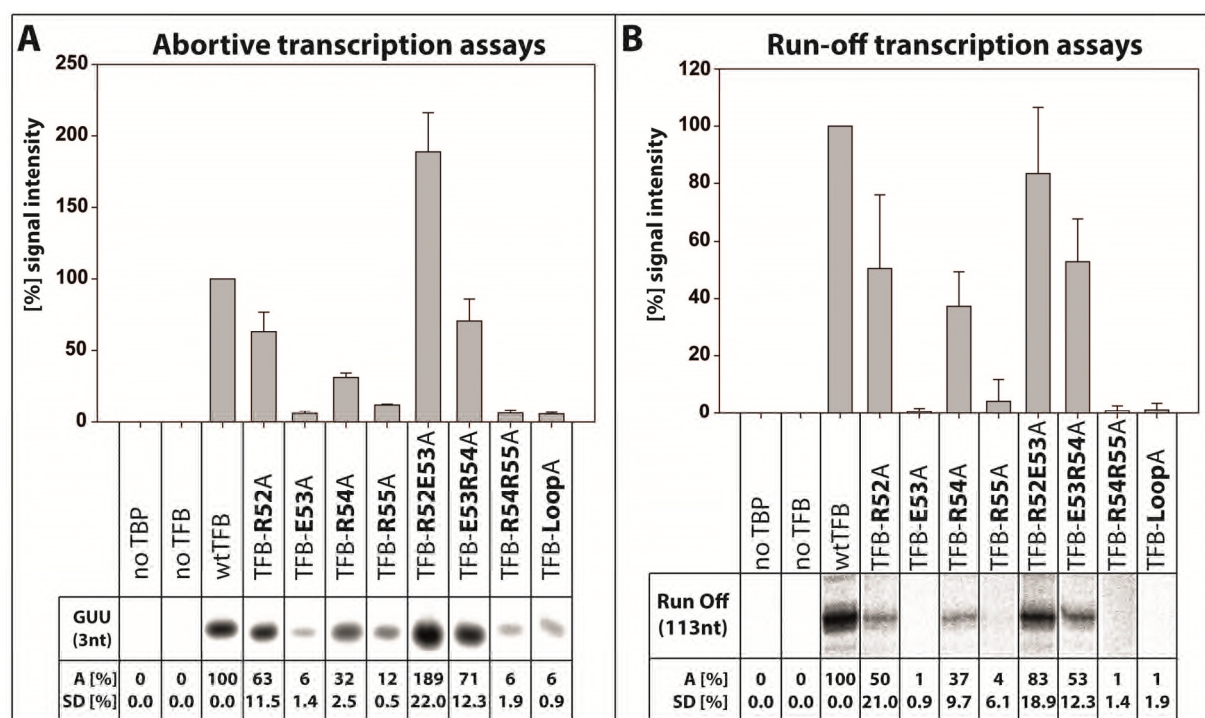
Figure 18: Results of abortive and run-off transcription assays with TFB alanine variants. **A)** Results of three independently performed abortive transcription assays are summarized in a bar diagram. Formation of the 3nt radiolabeled product occurs only in presence of all transcription components, TBP, wtTFB and RNAP (first two lanes). Standard deviation (SD) and average (A) of the quantified signal intensities of the respective TFB substitutions in comparison to wtTFB were calculated. **B)** Three independently performed run-off transcription assays with the respective TFB alanine substitutions are summarized in a bar diagram. Run-off transcripts are formed only in presence of transcription factors and RNAP (first two lanes). The standard deviation (SD) and the average (A) of the formed transcripts were calculated by comparison with the wtTFB signals and are given below.

for TFB-R52A and TFB-E53R54A, respectively. TFB-R54A had almost the same level of produced transcripts with a value of 37% in comparison to the 32% of the abortive transcription assay. As it was expected from initiation assays, the substitutions TFB-E53A, TFB-R55A, TFB-R54R55A, and the TFB-LoopA showed no signals on the gel, indicating that no transcripts were formed. TFB-E53R54A showed transcript formation at a level comparable to the wild type TFB. The stimulatory effect observed in the abortive transcription assay showed no impact in multiple-round transcription assays. The finding, that half of the used TFB variants showed no or massively reduced transcription levels in abortive and run-off assays, and the fact, that this region is located very closely to the transcribing strand in the initiation complex, leads to the assumption that the t-strand of the transcription bubble is not correctly stabilized.

2.   KMnO₄ footprint experiments of TFB B-reader alanine variants

To analyze the quality of promoter opening with the different TFB variants $KMnO_4$ footprint experiments were performed (Figure 19). A DNA template was used which is radiolabeled at the 5´end of the non-transcribed strand together with potassium permanganate. This reagent preferably cleaves single stranded DNA at T-residues, and therefore single stranded regions of the template can be visualized using a sequencing gel. Control samples were used to show the specificity of the formation of the initial single stranded transcription bubble together with a sequence ladder to determine the location of the open region of the template (Figure 19 A, B). The sequence from -11 to +8 relative to the transcription start site of the

non-transcribed strand is shown (Figure 19 A), whereas the initial transcription bubble has a distance from -9 to +5 (Spitalny, Thomm 2003). T-residues are located at -6, -4, -2, +2 and +3 within this region, therefore five distinct bands are present on the gel if both strands were separated and stabilized correctly. Control samples showed that if TBP, TFB or RNAP is
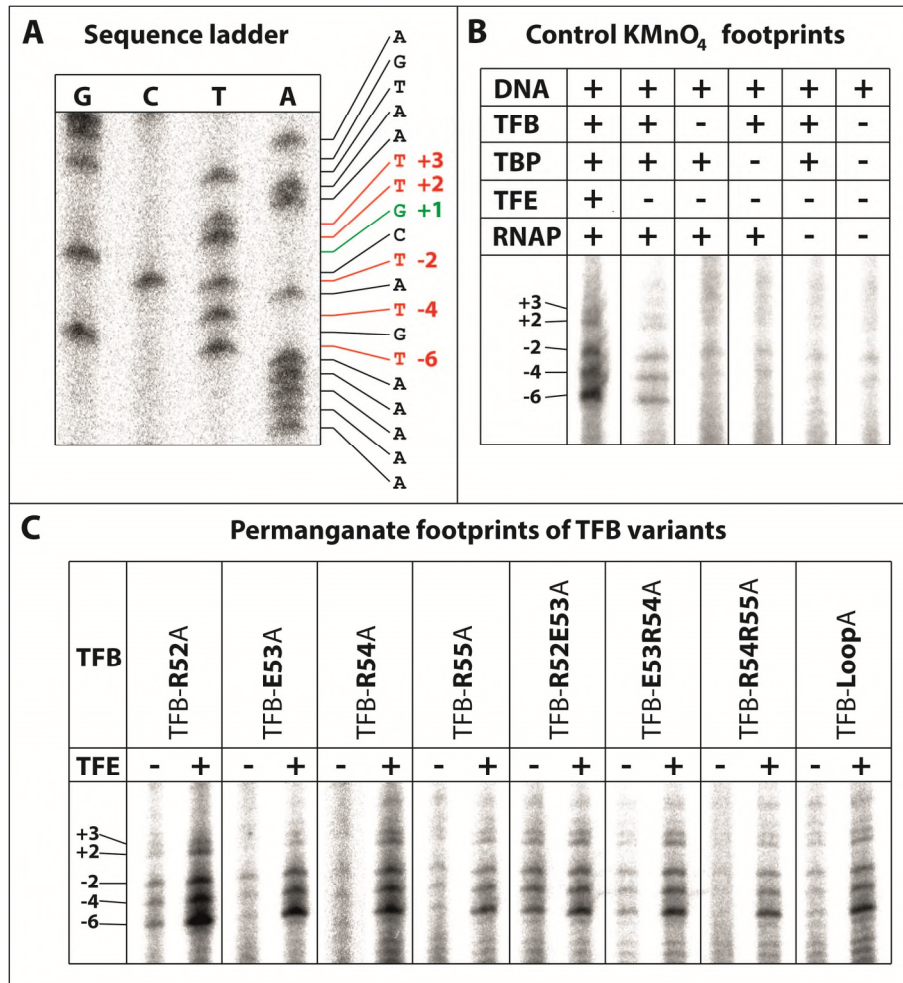

Figure 19: KMnO$_4$ footprint analysis of the TFB alanine variants. **A)** Sequence ladder of the non-transcribed strand of the initially melted DNA region of the used gdhC20 template. Sanger reactions were performed with dideoxynucleotides and radiolabeled M13R primer and separated on a 6% sequencing gel. TSS is shown in green (+1), and the T-residues within the initially melted region are colored in red. **B)** KMnO$_4$ footprints of controls using wtTFB. The distinct pattern consisting of five bands only occurred in presence of TFB, TBP, and RNAP, whereas additional TFE increases the signal intensity. **C)** Footprint analysis of the TFB variants was performed with and without TFE.

missing in samples with labeled DNA, no promoter opening takes place (Figure 19 B). However, if both transcription factors are present together with RNAP, the specific pattern can be observed, representing the single stranded region on the template DNA. Additional TFE increases the signal, as this transcription factor stabilizes the nt-strand by direct interaction with DNA (Grünberg et al. 2007). The analysis showed that TFB-E53A, TFB-R54A and the double substitution TFB-R54R55A are insufficient to open the promoter correctly, whereas addition of TFE can rescue the defects in promoter opening to some extent (Figure 19 C). TFB-R55A and TFB-LoopA showed a very low ability to open DNA, but addition of TFE leads to a weak opening of the DNA in case of both TFB variants. TFB-E53R54A showed the distinct pattern in both absence and presence of TFE, whereas the mutation TFB-R52A and TFB-R52E53A showed a good opening and a strong signal in

presence of TFE, indicating a proper melting and stabilization of the initial transcription bubble. However, almost all TFB variants had a weak signal of the +2 and +3 signals on the gel in absence of TFE, indicating that this partial region of the DNA can not be stabilized correctly during promoter opening.

3. TFE can partially compensate defects in promoter opening

It was shown that TFE can rescue defects in promoter opening due to a better stabilization of the single-stranded region of the transcription bubble in preinitiation complexes (Werner, Weinzierl 2005). Therefore abortive and run-off transcription assays were performed with additional TFE to see if some of the TFB variants can overcome defects in transcription bubble stabilization and form transcripts at a level comparable to the wtTFB (Figure 20). Assays were performed with and without TFE to compare the increase of the formed transcripts. The signal intensity of wtTFB showed an increase of 5.3-fold in abortive transcription assays if TFE is present (Figure 20 A). Additionally, a 5-7-fold increase was also observed for TFBR52A, TFB-R54A, TFB-R55A, and the two double-substituted proteins, TFB-R52E53A and TFB-E53R54A, in comparison to samples in which TFE was absent. Despite the similar activation fold in comparison to the wtTFB, the relative number of formed transcripts is almost constant between the above mentioned TFB variants and wtTFB in both, absence and presence of TFE. For example, TFB-R52A showed nearly 50% signal intensity relative to the wtTFB signal without TFE. The sample with additional TFE showed a signal of 60% relative to the wtTFB + TFE sample, whereas the increase of 7-fold was the highest observed for the TFB variants. Nevertheless, all TFB mutations mentioned before reached a level higher than wtTFB without TFE, except TFB-R55A. Therefore TFE can not rescue the defect of this mutation during transcription initiation. A more dramatic effect was observed for the mutations TFB-E53A, TFB-R54R55A and TFB-LoopA. Addition of TFE did not show any effects. Surely, there is also a slight increase of the formed transcripts with activation folds in the range of 2-3, but the fact, that these mutations do not show nearly any formation of transcripts, even in presence of TFE, it can be concluded that the substitution of the amino acids with alanine at these positions lead to a complete collapse of the transcription initiation. The results of these TFB variants in KMnO$_4$ footprint experiments indeed showed that the bubble is opened especially at the nt-strand in presence of TFE. Because of the location of the TFB B-reader loop a defect in the stabilization of the t-strand can not be excluded at this level. Proper nt-strand stabilization and insufficient t-strand stabilization can also lead to complexes unable to initiate transcription. However, to investigate if the loop is involved in strand separation, a run-off transcription experiment was performed in presence of TFE (Figure 20 B). In this assay RNA has to pass the loop correctly, otherwise RNA is not guided towards the exit channel, and a run-off signal should not be observed. The wtTFB showed an increased signal intensity of 2.2-fold in presence of TFE in comparison to the sample without TFE. TFB-R52A, TFB-R52E53A and TFB-E53R54A showed increased signal intensities of 1.8 - 2.7-fold, which is a comparable level to wtTFB. In comparison to abortive transcription assays, this result was expected, as these TFB variants showed also similar activation folds relative to the wtTFB. Interestingly, TFB-R55A, which had a very weak signal in the abortive transcription assay without TFE, does also not show a run-off product in multiple round transcription assays. But addition of TFE can raise the signal above the wtTFB niveau and further showed transcripts of 50% in comparison to wtTFB + TFE sample. In contrast, TFB-E53A, TFB-R54R55A and the TFB-LoopA showed no formation of the run-off product in absence of TFE, but addition of this transcription factor led to a slight run-off signal of 13% and 22%. TFB-E53A showed no run-off product even in presence of TFE, which indicate that
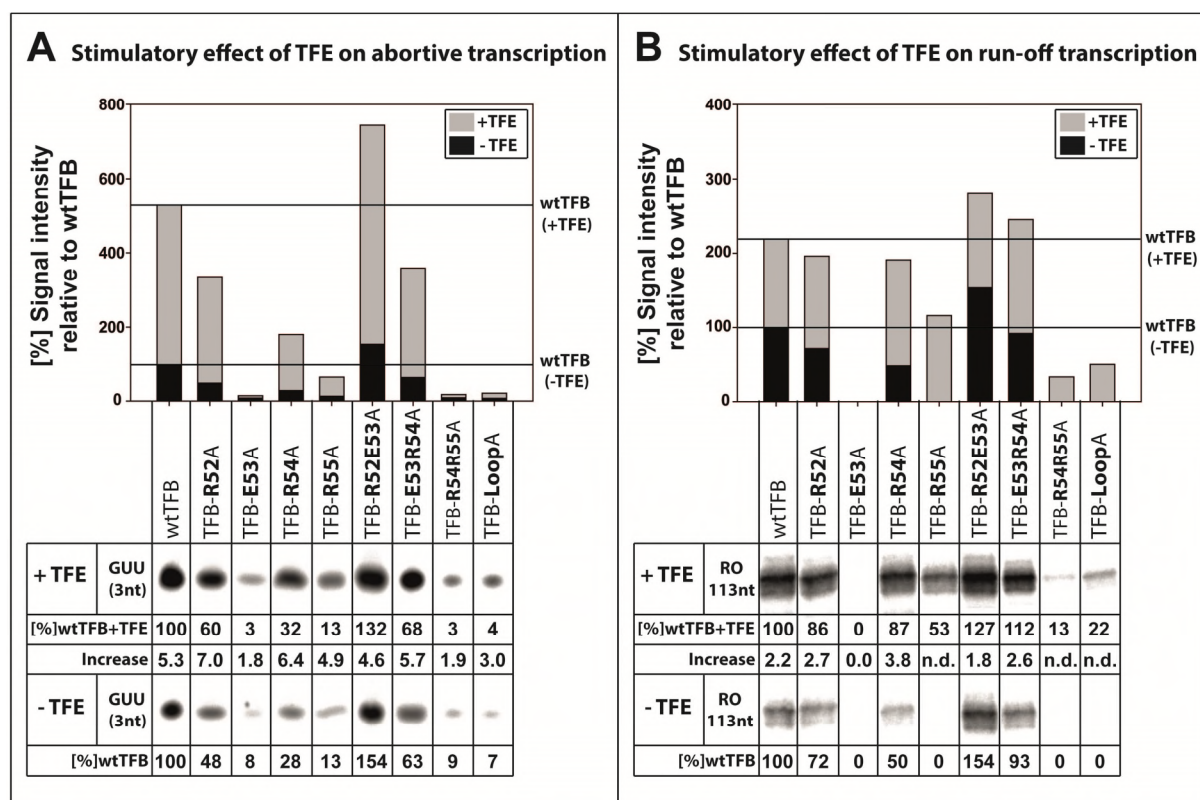
Figure 20: Abortive and run-off transcription assays in absence and presence of TFE. **A)** Results of abortive transcription without TFE (-TFE) and with TFE (+TFE) are summarized in a bar diagram. Black bars represent the quantified signal intensity of the radiolabeled 3nt RNA product of the TFB proteins without TFE relative to the quantified signal intensity of the 3nt RNA derived from wtTFB without TFE. Grey bars represent the signal intensity of RNAs quantified from samples containing TFE in comparison to wtTFB + TFE. In both cases (-TFE and +TFE) the respective percentages were calculated and given below the signals. The activation fold (increase) of samples containing TFE was calculated by comparison with the signal intensity of samples without TFE, whereas wtTFB niveaus are depicted as horizontal bars at 100% (-TFE) and 530% (+TFE). **B)** Run-off transcription assays in absence and presence of TFE. The depiction is the same as in A. Horizontal lines at 100% and 220% represent the wtTFB signal with and without TFE, respectively. (n.d. = not defined)

a negative charge at this position of TFB is essential for transcription. The results of the run-off transcription assay together with the results of the abortive transcription assay showed that some of the TFB variants can not initiate transcription even in presence of TFE. This finding leads to the suggestion that the t-strand is not correctly stabilized. However, to investigate the RNA-strand separation in particular, the used transcription assays are insufficient to give answers to this question, because some TFB variants showed massive defects in transcription assays. To overcome this problem a pre-opened template was used which had a mismatch at registers -1 to +2 and therefore contained a mini-bubble around the start site.

4. RNA-strand separation at heteroduplex DNA templates

To see if the TFB-A variants are involved in the RNA-strand separation, a DNA template containing a mini-bubble was used. Two templates were generated, whereby one contained the regular t-strand sequence and a mismatch sequence at the nt-strand, and vice versa (Figure 21 A). To verify if the pre-opened region is large enough to enable nearly wild type transcription levels of the TFB variants, an abortive transcription assay was performed (Figure 21 B).

This experiment showed that almost all TFB alanine substitutions had wtTFB levels in formation of the 3nt RNA transcript, except TFB-E53A (31%), TFB-R54R55A (66%) and TFB-LoopA (83%). The hybridized control template which lacks the mismatch region showed a 20-fold decreased activity, indicating that the pre-opening of the initially melted region
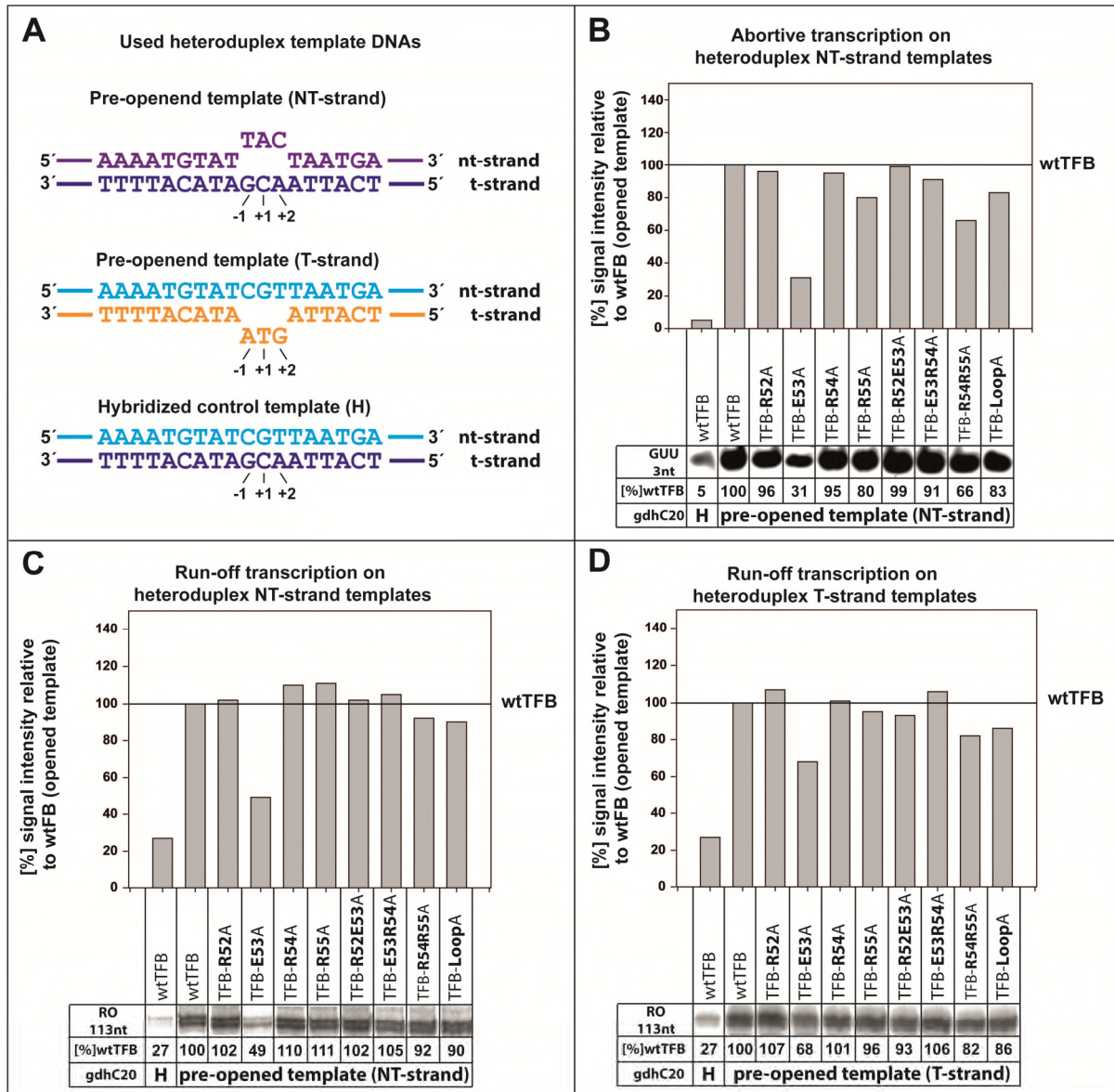
Figure 21: Transcription assays using pre-opened templates. **A)** Schematic representation of the used templates. The pre-opened nt-strand template contained a mismatch at registers -1 to +2 at the nt-strand (purple) whereas the t-strand had the regular sequence (dark blue). The pre-opened t-strand template contained an nt-strand with the regular sequence (light blue) and a t-strand mismatch at registers -1 to +2 (orange). The control template was hybridized in the same manner like the pre-opened templates, but contained the regular sequences of both the t-strand (dark blue) and the nt-strand (light blue). **B)** Abortive transcription assay with a template containing a mismatch at the nt-strand. Signals were quantified and compared with the signal intensity of the wtTFB sample in which the heteroduplex template was used. The relative signal intensities were summarized as bar diagrams and the wtTFB level is represented as a horizontal black line. Assays performed with the hybridized control template (H) as well as absolute values of the quantification are given below the signals. **C)** Run-off experiment with a template containing a mismatch at the non-transcribed strand. Depiction is the same as in B. **D)** Run-off transcription experiment with a template containing a mismatch at the transcribed strand. Depiction is the same as in B and C.

55

from registers -1 to +2 supports the initial synthesis of RNA in the abortive transcription assay. The reduction of the signal intensity of TFB-E53A of 31% further showed that also a pre-opening of the template does not support formation of the 3nt RNAs in sufficient amounts, underlining the importance of this amino acid position. However, the same assay with the pre-opened t-strand could not be performed, as the mismatch creates a region where a matching dinucleotide primer was not available. Nevertheless, both templates were used in run-off transcription assays to investigate if the charge of the B-reader loop region is involved in RNA strand-separation. The run-off transcription assay with the mismatch at the nt-strand showed similar results like in the respective abortive transcription assay (Figure 21 C). First, the signal intensity of wtTFB raised about 3 - 4-fold in comparison to the hybridized control template, indicating that pre-opening of DNA supports transcription. In contrast to the control template, two distinct bands were located at the respective site on the gel at 113nt which is the run-off length and a second band below, probably at 112 or 111nt. This second band only occurs in presence of the pre-opened template, suggesting that use of this template allows non-selective diffusion of NTPs to the start site, whereas transcription initiates at two different start points, +1 and +2. However, almost all TFB variants showed transcripts comparable to the wtTFB level, indicating that the use of a heteroduplex DNA template can overcome the deficiencies in transcription bubble stabilization observed in previous experiments on the one hand. It further shows that the amino acids substituted with alanine at the respective sites seem to be not involved in RNA-strand separation on the other hand. A defect of the separation of the RNA can be excluded, because RNA was guided correctly towards the RNA exit channel, leading to the formation of run-off transcripts at a level comparable to wtTFB. The only TFB substitution with an altered signal intensity was TFB-E53A with only 50% of the formed transcripts in comparison to the wild type TFB. This finding shows that a pre-opening of the template at the nt-strand leads to more transcripts in comparison to the experiment with additional TFE, but the reduced amount of formed transcripts further indicate that the amino acid E53 possibly plays a key role in t-strand stabilization and/or RNA-separation. To investigate the impact of the E53A substitution on the t-strand stabilization, a run-off transcription assay with a mismatch at the t-strand was performed (Figure 21 D). The wtTFB also showed an increase of the formed transcripts of 3 - 4-fold in comparison to the regular, PCR-produced template and the hybridized control template. The other TFB variants tested showed also transcription levels comparable to the wtTFB, indicating that a mismatch at the t-strand supports transcription in a sufficient way. Only TFB-R54R55A and TFB-LoopA showed slightly decreased values in comparison to the experiment in which the nt-strand mismatch was used. TFB-E53A formed transcripts of 68% in comparison to the wtTFB and therefore a mismatch at the t-strand increase the transcription rate of this TFB substitution in comparison to the mismatch at the nt-strand. This indicates that the amino acid E53 interacts with the t-strand preferably to stabilize the DNA in the open complex. It might also support RNA strand separation, but if this negative charged amino acid is sufficient for the separation only, the number of transcripts should be less than the observed 68%. However, the results of the experiments using mini-bubbles as templates showed that the transcription output is almost at the wtTFB level, demonstrating that the charge of the B-reader loop is not important for the RNA-strand separation. Moreover, the B-reader loop tip region is sufficient for the stabilization of the t-strand in open complexes, whereas the only negative charged amino acid E53 plays a key role during transcription initiation, possibly due to charge-dependent interactions with the transcribing strand.

5. Summary of the TFB alanine substitutions

The TFB B-reader loop tip region was substituted with alanine amino acids to stepwise eliminate the overall positive charge of this region to investigate the role of this protein domain during transcription initiation with respect to charge-dependent interactions between the loop and nascent RNA. Different *in vitro* transcription assays were performed to observe impacts of the TFB variants during transcription, and are summarized in figure 22. The results presented here demonstrate that the variants TFB-R52A, TFB-R54A and TFB-E53R54A can form preinitiation complexes, initiate transcription and form run-off products similar or slightly reduced in comparison to the wtTFB. Experiments with additional TFE lead also to transcription levels comparable to the wild type niveau. This finding indicates that the charge of the regular amino acids at these positions is not important for the function of the

TFB variant	loop tip				PIC formation	OC formation (-TFE)	OC formation (+TFE)	first bond formation	run-off transcription	first bond formation	run-off transcription	first bond formation (nt-strand mismatch)	run-off transcription (t-strand mismatch)	run-off transcription (nt-strand mismatch)
										TFE rescue		**Pre-opened templates**		
wtTFB	R	E	R	R	++	++	++	++	++	++	++	++	++	++
TFB-R52A	A	E	R	R	++	++	+++	+	+	+	++	++	++	++
TFB-E53A	R	A	R	R	+	-	+	-	-	-	-	+	++	+
TFB-R54A	R	E	A	R	++	-	++	+	+	+	++	++	++	++
TFB-R55A	R	E	R	A	++	-	+	-	-	-	+	++	++	++
TFB-R52E53A	A	A	R	R	++	+++	++	+++	++	+++	++	++	++	++
TFB-E53R54A	R	A	A	R	++	+	++	+	+	+	++	++	++	++
TFB-R54R55A	R	E	A	A	++	-	+	-	-	-	-	+	++	++
TFB-LoopA	A	A	A	A	++	-	++	-	-	-	-	++	++	++

■ unpolar, hydrophobic ■ positive charge ■ negative charge
- = insufficient   + = weak   ++ = moderate, wt-level   +++ = increase

Figure 22: Summary of the experiments performed with TFB alanine substitutions. All TFB variants are listed below the wtTFB, and the respective amino acid composition of the B-reader loop tip region is shown and colored with respect to the particular charge of the amino acids: unpolar (grey), basic (blue) and acidic (red). The results of the TFB variants are summarized below the respective experiment, whereas the quality is given as (-) for insufficient results, (+) for weak, (++) for moderate results or results comparable to wtTFB, and (+++) for increased values in comparison to the wtTFB.

B-reader loop region. In contrast, the double-substitution TFB-R52E53A showed increased initiation ability and a better opening of the nt-strand in $KMnO_4$ footprint experiments, indicating that this amino acid composition can stimulate transcription initiation, whereas in run-off experiments no significant difference in comparison to the wtTFB was observed. The most dramatic effects were observed for the TFB substitutions TFB-E53A, TFB-R55A, TFB-R54R55A and the TFB-LoopA. All variants formed preinitiation complexes, indicating that the integrity of the TFB proteins is unchanged and are sufficient to interact correctly with DNA, TBP and RNAP. However, none of the substituted proteins is able to initiate transcription, and no formation of run-off transcripts was observed. Addition of TFE leads to a better

opening of the initially melted region especially for the nt-strand in $KMnO_4$ footprint experiments, but in abortive and run-off transcriptions TFE can not rescue the defects, because the observed signal intensities of the respective experiments were below the wtTFB -TFE niveau. The experiments with mini-bubbles showed that the transcription level can nearly be equalized for the TFB variants, allowing investigation of the strand separation. The assays demonstrated that the charge of the B-reader loop tip region is not important for the separation of the RNA strand, because run-off transcripts were formed at a level comparable to wtTFB, indicating that RNA is guided correctly towards the RNA exit channel. Only TFB-E53A showed also low transcription levels in the assay with the nt-strand mismatch, and moderate levels in the assay with the t-strand mismatch. Taken together, the B-reader loop tip region is important for the stabilization of the t-strand of the transcription bubble in preinitiation complexes. The charge of this protein region is not important for the RNA-strand separation, whereas the only negative amino acid, E53, is essential for transcription.

D.   TFB-DNA crosslink studies during transcription initiation

The main chapter of this thesis is about structural rearrangements of the transcription factor B during transcription initiation and transition to early elongation. As it was mentioned in the outline of the previous chapter, a crystal structure of an initially transcribing complex of the eukaryotic *S. cerevisiae* organism was published and served as starting point of this thesis. The structure contains TFIIB, RNAP II, DNA and a 6nt RNA bound to the transcribed DNA strand (Figure 6). Several postulations derived from this structure, whereas one issue, the RNA-strand separation was addressed and discussed in the previous chapter. Beside the possible role of the TFB B-reader loop domain during transcription, it was also hypothesized that the B-reader domain is located in close proximity to the active site of the RNAP where it interacts with nascent RNA. A clash between the nascent RNA and the TFIIB B-reader helix was postulated to guide the transcript towards the exit channel of the polymerase. This interaction should take place at an RNA length of 8nt, as it was suggested in a published open complex model (Kostrewa et al. 2009). A further clash of the RNA with the Zn-ribbon domain of the transcription factor IIB was further postulated at an RNA length of 12-13nt because this protein region blocks the exit pore of the channel (Sainsbury et al. 2013).

No structural information is available for the *Pyrococcus* transcription system with respect to the topology of TFB within the complex, and structural rearrangements and possible interactions with RNA and DNA during transcription initiation. To investigate the archaeal TFB of *P. furiosus* (*Pfu*TFB) during transcription, a UV-inducible crosslinking system was used in this study. Based on the relationship between archaeal and eukaryotic transcription machineries, corresponding amino acids of the *Pfu*TFB were substituted with p-Benzoyl phenylalanine (Bpa). This phenylalanine derivate reacts preferentially with unreactive C-H bonds when exposed to UV-light at a wavelength of 350-360nm with a reactive spherical radius of 3.1Å (Kauer et al. 1986; Dorman, Prestwich 1994). If TFB-Bpa variants were incubated with RNAP, TBP and DNA and exposed to UV-light, it forms a covalent crosslink with nucleotides in proximity to the Bpa position. To visualize this interaction, radioactively labeled nucleotides were used at specific sites of the DNA, whereas TFB-Bpa/DNA complexes were analyzed by SDS-PAGE. The principle of this method is shown in figure 23. This approach enables investigation of TFB-DNA contacts in preinitiation complexes and initially transcribing complexes which in turn allows comprehending structural transitions and functional interactions of TFB during transcription initiation and transition to elongation. The presented data shown here are the first biochemical data on dynamic transitions of the archaeal transcription factor B during transcription initiation and transition from initiation to early elongation.

The first step was the selection of TFB-Bpa variants suitable for crosslinking experiments. A set of TFB mutants were created including amino acid positions G41 - A49 of the *Pfu*TFB B-reader helix (corresponding positions R60 - N68 of *Saccharomyces cerevisiae* (*Sc*) TFIIB), and S50 - R57 of the *Pfu*TFB B-reader loop (corresponding positions D69 - R78 of *Sc*TFIIB). Both domains are located closely to the active site of the RNA polymerase in the initially transcribing complex (Figure 29). In addition, *Pfu*TFB linker positions E74 and M85 (R95 and M104 of *Sc*TFIIB) as well as the *Pfu*TFB B-core position F192 (I209 of *Sc*TFIIB) were substituted with Bpa for additional crosslink reactions and controls. Because of the potential impact of unnatural amino acids on protein stability and function the TFB variants were screened in a set of *in vitro* transcription reactions.
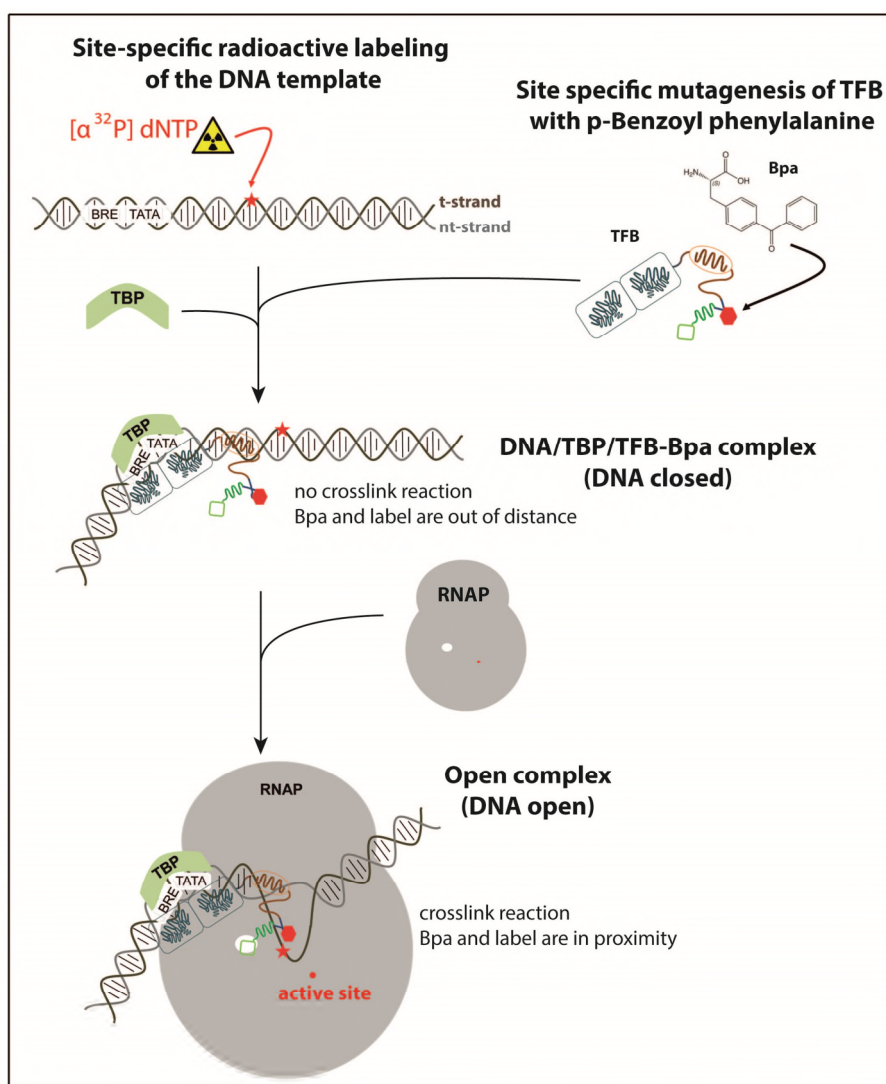
Figure 23: Schematic draw of the used crosslinking method. DNA-templates containing a strong glutamate dehydrogenase (gdh) promoter are radioactively labeled at one specific site using [α-^{32}P] dNTPs. Selected amino acids of the *Pfu*TFB were substituted with the UV-inducible photo-reactive p-Benzoyl phenylalanine (Bpa). Both were incubated with TBP and/or additional RNAP to form preinitiation complexes. When RNAP is absent, a DNA/TBP/ TFB-Bpa complex is formed, whereas the DNA is double stranded (closed DNA). The label at the DNA is not within the reactive radius of the Bpa, and no crosslink reaction can be observed. Addition of RNAP leads to the formation of open complexes, whereas Bpa and the label are in close proximity to form a covalent TFB-DNA bond.

1.   Analysis and selection of TFB-Bpa variants

First, the ability of TFB-Bpa proteins to form a preinitiation complex was analyzed using EMSAs (Figure 24). Control samples were performed with a 5´end FAM labeled gdh-C20 template in absence of proteins (Figure 24 A, lane 1), in absence of TFB (lane 2), and with TFB only (lane 3). A specific signal on the gel occurred only in presence of DNA, TBP and wtTFB, indicating the ternary DNA/TBP/TFB complex (lane 4). Addition of RNAP shifted the signal to a band located higher on the gel, indicating preinitiation complex formation (lane 5). The TFB-Bpa variants were tested in the same manner as wtTFB, whereas the TFB-linker positions TFB-E74Bpa and TFB-M85Bpa, as well as the B-core TFB-F192Bpa showed signal patterns comparable to the wtTFB. The TFB B-reader helix variants G41Bpa - A49Bpa can also form PICs (Figure 24 B, lanes 4 - 11) whereas the TFB-P42Bpa was indeed successfully created on the plasmid, but overexpression failed several times for unknown

reasons and is therefore missing in the screening. The TFB B-reader loop mutations S50Bpa - R57Bpa showed also similar patterns (Figure 24 B, lanes 12 - 19), indicating the presence of PICs. Therefore, all TFB variants tested have the ability to form stable complexes in presence of DNA, TBP and RNAP, indicating that the incorporated unnatural amino acids do not affect PIC formation.
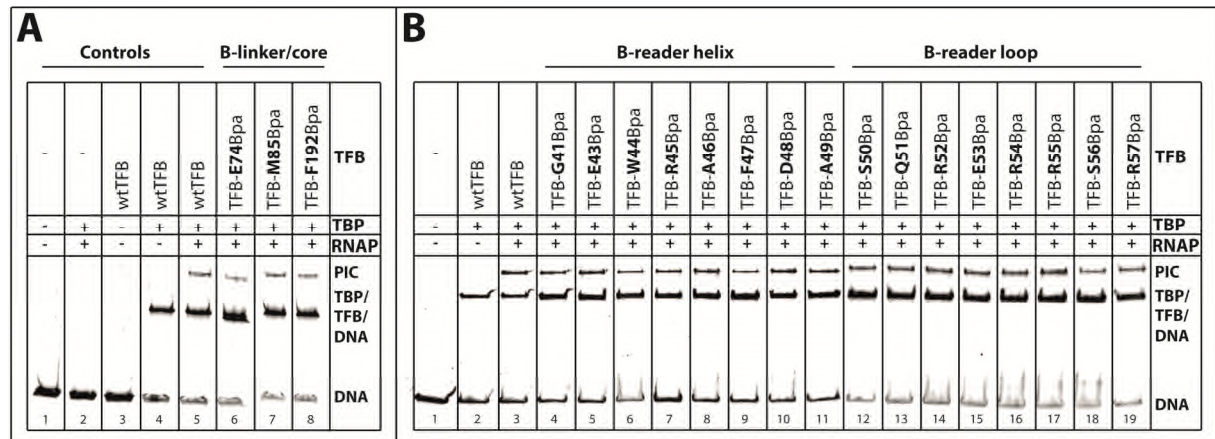
Figure 24: Electro mobility shift assays of TFB-Bpa variants showed preinitiation complex formation. **A)** Control samples with DNA (lane 1), DNA, TBP and RNAP (lane 2) and DNA with TFB (lane 3) showed no unspecific interaction. Reactions containing DNA, TBP and wtTFB (lane 4) formed ternary DNA/TBP/TFB complexes indicated by the shift of the signal, whereas in presence of RNAP an additional band is visible, indicating the preinitiation complex (lane 5). The B-linker positions E74Bpa (lane 6) and M85Bpa (lane 7), as well as the B-core position F192Bpa (lane 8) formed preinitiation complexes. A native 4-10% gradient gel was used. **B)** Control samples were performed with wtTFB (lane 1-3). All TFB-Bpa positions of the B-reader helix domain (G41Bpa - A49Bpa; lane 4 - 11) formed PICs comparable to the wtTFB. The B-reader loop positions S50Bpa - R57Bpa (lane 12 - 19) showed similar patterns in comparison to wtTFB (lane 3), indicating PIC formation. A native 4-20% gradient gel was used.

To test the ability of the formation of the first phosphodiester bond of complexes initiated with TFB-Bpa variants, abortive transcription assays were performed (Figure 25). The results of experiments performed with the TFB B-linker positions E74Bpa and M85Bpa showed that the signal intensities of the labeled 3nt RNAs had the same signal intensity as signals of the wtTFB, indicating that the incorporated unnatural amino acid at these B-linker sites have no influence on first phosphodiester bond formation (Figure 25 A). In contrast, the B-core mutation F192Bpa showed transcription activity of two-third in comparison to the wtTFB (Figure 25 A). This amino acid position was chosen because of its close location to DNA (Kosa et al. 1997), and the incorporated Bpa may influence the binding of TFB-F192Bpa to DNA/TBP which leads to reduced signal intensities in abortive transcription assays. The results of the TFB B-reader helix Bpa variants are summarized in figure 25 B. Substitutions G41Bpa, E43Bpa, R45Bpa and D48Bpa showed formation of 3nt RNAs between 10% - 20% in comparison to the wtTFB. This finding suggests that the substitution of the natural amino acids at these positions with Bpa influences transcription in a way that open complex formation, transcription start site selection or other important interactions within the complex are insufficient to successfully form the first phosphodiester bond. TFB-F47Bpa and TFB-W44Bpa showed reduced initiation levels in comparison to the wtTFB with values of 28% and 38%, respectively. Only TFB-A46Bpa and TFB-A49Bpa showed moderate transcript levels of nearly 60-65% in comparison to the wild type. The screening of the TFB B-reader loop variants also showed that some TFB substitutions are not able to initiate the synthesis
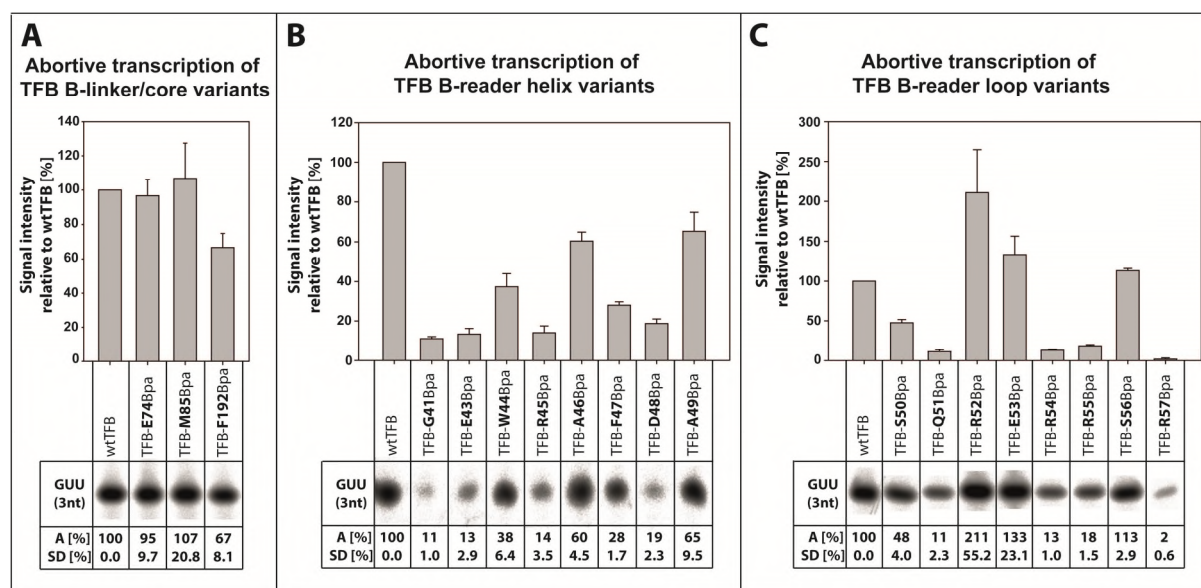
Figure 25: Abortive transcription assays of TFB-Bpa variants. **A)** Results of the B-linker positions E74Bpa and M85Bpa together with the B-core mutation F192Bpa are summarized as a bar diagram in comparison to wtTFB. Signals derived from formation of the 3nt radiolabeled RNAs were quantified and compared to the signal intensity of wtTFB. Reactions were performed in three independent experiments, whereas the average (A) and the standard deviation (SD) were calculated and shown below the signals. **B)** TFB B-reader helix Bpa variants are shown in comparison to the wtTFB. Depiction is the same as in A. **C)** TFB B-reader loop Bpa variants are shown in comparison to wtTFB. Depiction is the same as in A.

of the 3nt RNA product (Figure 25 C). TFB-Q51Bpa, TFB-R54Bpa and TFB-R55Bpa showed values of the quantified transcripts between 11% - 18%, whereas TFB-R57Bpa is not able to initiate RNA synthesis. TFB-S50Bpa showed bisected values relative to the wtTFB signals. In contrast, TFB-E53Bpa and TFB-S56Bpa had slightly increased values in the abortive transcription assays with values of 133% and 113% relative to the wtTFB, indicating that the substitution with Bpa at these positions does not influence transcription. TFB-R52Bpa showed the highest value of the quantified transcripts with 211% in average. Therefore Bpa stimulates abortive transcription if substituted with the natural arginine at this position. In addition to the abortive transcription assays run-off transcription assays were performed to analyze TFB-Bpa variants in multiple transcription rounds (Figure 26). Here, similar results were observed in comparison to the initiation assays. The TFB B-linker positions E74Bpa and M85Bpa again showed transcripts at a level comparable to the wtTFB (Figure 26 A). The B-core position F192Bpa also showed signal intensities of the quantified run-off product with a value of 95% in average, indicating that this mutation does not hamper the transcription recycling process over time and therefore only abortive transcription is influenced by this mutation. Due to the weak ability for transcription initiation, the B-reader helix mutations G41Bpa, E43Bpa, R45Bpa and D48Bpa showed no formation of the run-off product, as it was expected (Figure 26 B). Also W44Bpa and F47Bpa showed decreased values in these assays around 20%, whereas the A46Bpa and A49Bpa had transcription levels with averaged values of 40% and 48%, respectively. Run-off transcriptions with the TFB B-reader loop substitutions showed that position Q51Bpa and R57Bpa could not form transcripts (Figure 27 C), whereas R54Bpa and R55Bpa showed strongly decreased transcription about 27-34% relative to the wtTFB. Only S50Bpa, E53Bpa and S56Bpa showed moderate transcription levels comparable to the wtTFB with averaged values of 62%, 68% and 75%, respectively, whereas R52Bpa formed run-off transcripts comparable to wtTFB (119%).
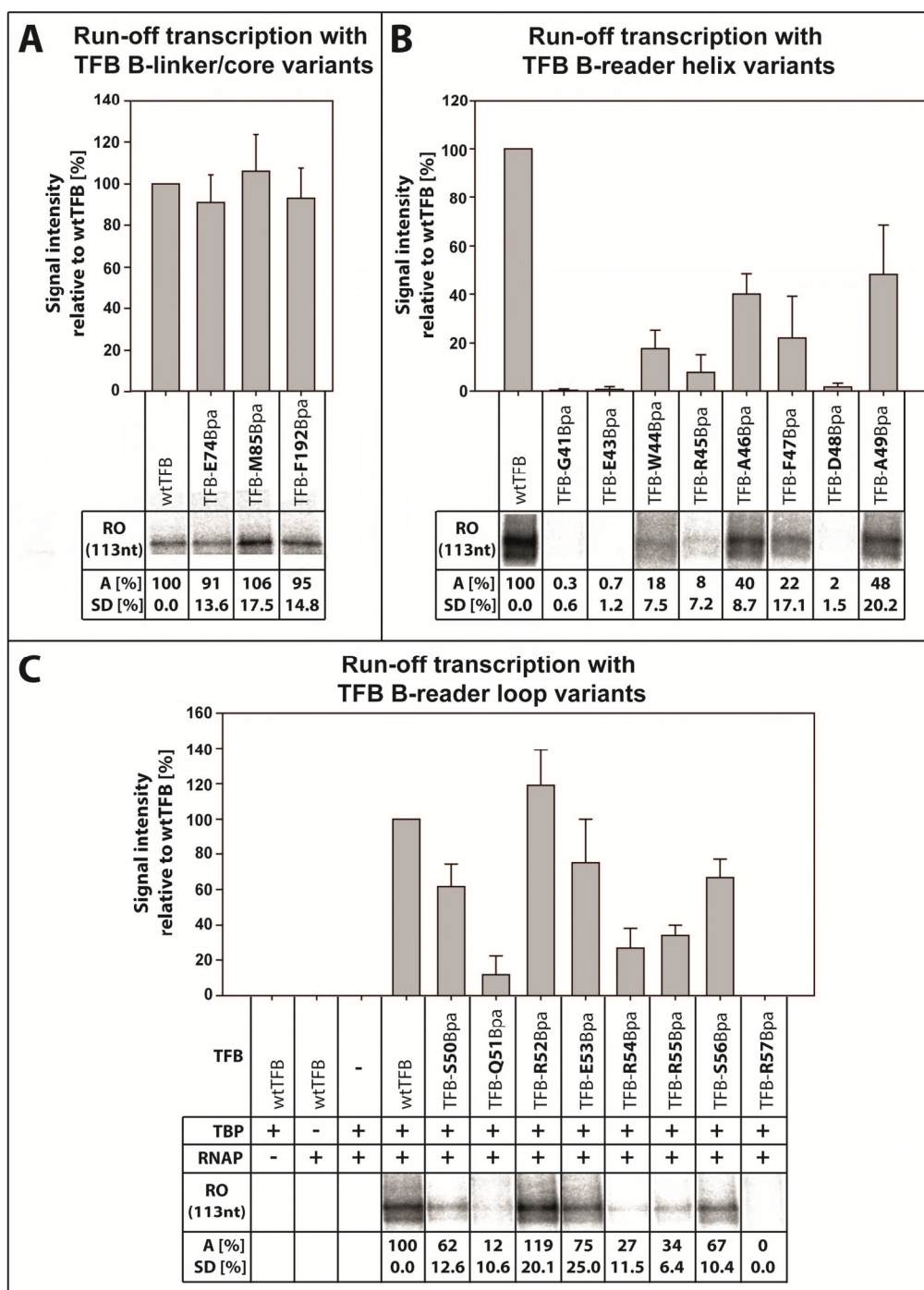
Figure 26: Summary of the run-off transcription experiments with TFB-Bpa variants. **A)** Three independently performed transcription assays were summarized as a bar diagram for the TFB B-linker positions E74Bpa and M85Bpa, and the B-core substitution F192Bpa. Signals derived from the 113nt radiolabeled RNA transcripts were quantified and compared to the signals of the wtTFB. The average (A) and the standard deviation (SD) are shown below the signals. **B)** Experiments of the TFB B-reader helix variants. Depiction is the same like in A. **C)** Control assays were performed in absence of RNAP, TBP and TFB to show that the formation of the 113nt RNA depends on the presence of all transcription factors. The B-reader loop Bpa variants are summarized as bar diagrams whereas the depiction is the same as in A.

To analyze if the amino acid substitutions affect the promoter opening event, a KMnO$_4$ footprint experiment was performed (Figure 27). Here, it was investigated if the defects of the TFB-Bpa variants, which showed very weak or no activity in transcription, can be rescued by the addition of TFE as it was pointed out by (Werner, Weinzierl 2005). Control assays and a sequence ladder are shown in the previous chapter IV. C. 2 Figure 19, whereas only samples performed in presence of TFE were shown to analyze defects in promoter opening which can not be compensated by TFE. The results showed that TFB-mutations G41Bpa, D48Bpa, Q51Bpa and R57Bpa are not able to open the initially melted region around the transcription start site in presence of TFE, which would explain the results of the transcription assays. TFB variants E43Bpa, R45Bpa, F47Bpa, R54Bpa and R55Bpa showed weak opening of the DNA in presence of TFE, whereas W44Bpa, A46Bpa, A49Bpa, S50Bpa, S56Bpa, E74Bpa, M85Bpa and F192Bpa showed signals on the gel comparable to the wtTFB. Only R52Bpa and E53Bpa showed increased ability to open the DNA. The Footprint experiment demonstrated that substitutions of the natural amino acids at distinct positions with Bpa can negatively influence the promoter opening, resulting in a very weak or lost transcriptional activity.

Analysis of the used TFB-Bpa variants showed that indeed all proteins are able to form preinitiation complexes, but the ability to initiate transcription is restricted to a few Bpa positions only. As it turned out in footprint experiments, incorporation of the unnatural amino acid can inhibit promoter opening for some TFB-Bpa mutations. Therefore transcription initiation and run-off transcription is weak. Based on this analysis only TFB variants were selected which showed applicable transcription levels.
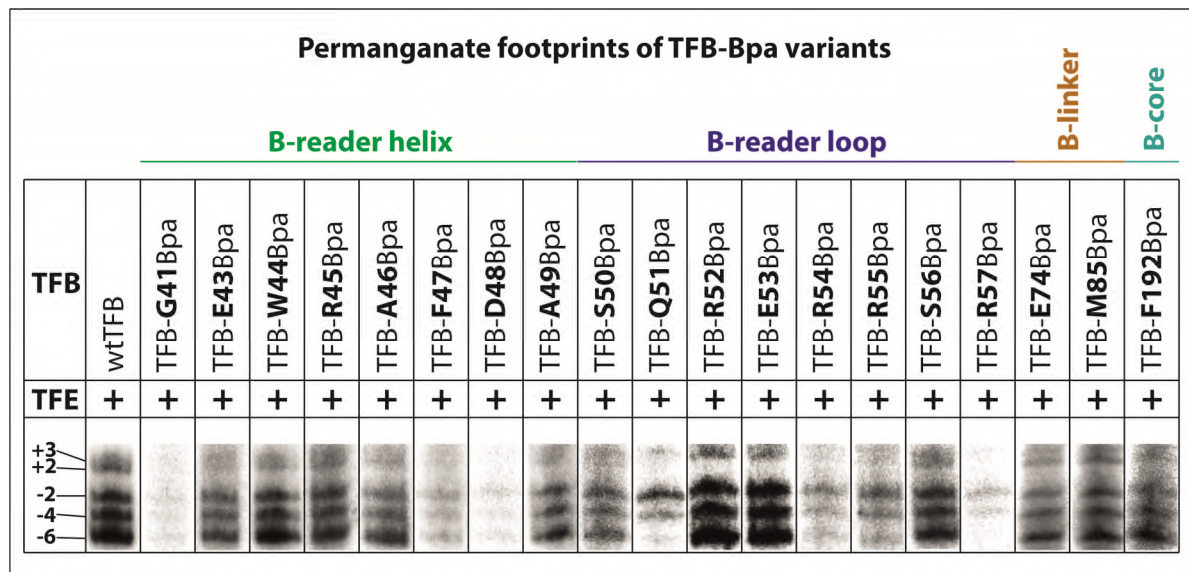
Figure 27: Footprint analysis of the used TFB-Bpa variants with respect to TFE compensation. A gdh-C20 template was used, whereas control assays and the sequence ladder are shown in figure 19 (Chapter IV. C. 2). The specific patterns of the cleaved radiolabeled nt-strands are shown for each TFB-Bpa position in comparison to the wtTFB, whereas only samples containing TFE were depicted. Domains of the TFB are indicated as horizontal bars and colored for the TFB B-reader helix (green), the B-reader loop (blue), the B-linker (brown) and the B-core (cyan).

The summary of the analyzed TFB-Bpa variants showed that in total 9 of 19 mutants tested were suitable for crosslink reactions (Figure 28). However, good results of the variants in the preliminary *in vitro* experiments were one criterion, whereas the topology of the amino acids was another important point of the selection. Therefore a set of six TFB variants were selected (Figure 29). TFB-A46Bpa was chosen because of its location in the helical domain of the B-reader, whereas TFB-A49Bpa and S50Bpa were not used in this study


	B-reader helix								B-reader loop								B-linker		B-core
	TFB-**G41**Bpa	TFB-**E43**Bpa	TFB-**W44**Bpa	TFB-**R45**Bpa	TFB-**A46**Bpa	TFB-**F47**Bpa	TFB-**D48**Bpa	TFB-**A49**Bpa	TFB-**S50**Bpa	TFB-**Q51**Bpa	TFB-**R52**Bpa	TFB-**E53**Bpa	TFB-**R54**Bpa	TFB-**R55**Bpa	TFB-**S56**Bpa	TFB-**R57**Bpa	TFB-**E74**Bpa	TFB-**M85**Bpa	TFB-**F192**Bpa
PIC formation	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++
OC formation	-	+	++	+	++	+	-	++	++	-	+++	+++	+	+	++	-	++	++	++
*Abortive transcription	11	13	38	14	60	28	19	65	48	11	211	133	13	18	113	2	95	107	67
*Run off transcription	-	-	18	8	40	22	-	48	62	12	119	75	27	34	67	-	91	106	95

= insufficient = weak = applicable = selected

- = insufficient + = weak ++ = wt level +++ = increase

∗ = Average [%] relative to wtTFB

Figure 28: Summary of the analysis of the tested TFB-Bpa variants. The results of each experiment are shown for every of the TFB-Bpa proteins. The ability to form a preinitiation complex tested in EMSAs as well as the quality of the promoter opening tested in permanganate footprints are indicated by (-) for insufficient, (+) weak, (++) moderate or wild type level and (+++) increased rates in comparison to the wild type TFB. For abortive and run-off transcription assays the averaged values of the respective experiments are shown in percentages relative to the wtTFB. The overall quality of the single TFB-Bpa positions are highlighted in red for insufficient, yellow for moderate and green for applicable results. The selected Bpa positions are highlighted in grey, whereas the corresponding TFB domain is shown above the amino acid positions and colored in green (B-reader helix), blue (B-reader loop), brown (B-linker) and cyan (B-core).

although they showed also good results in the analysis. To investigate the transition of the loop and its postulated interactions with nascent RNA, the TFB-S56Bpa was also chosen as this amino acid is located at the tip of the loop. The other suitable mutation of the loop region, TFB-E53Bpa, was not selected, but the amino acid position R52 next to it. TFB-R52Bpa had the best results observed in the analysis of the TFB variants, and was used to constitute and adjust the crosslinking setup. It is further part of the B-reader loop with the closest distance to the DNA in the crystal structure (Figure 29). At least, the two B-linker positions E74Bpa and M85Bpa, and the B-core position F192Bpa showed results comparable to the wtTFB and completed the set of the used TFB variants. These positions served as control positions to determine the crosslink specificity on the one hand, and they were used in crosslink reactions to determine the topology of the TFB within the complex on the other hand.
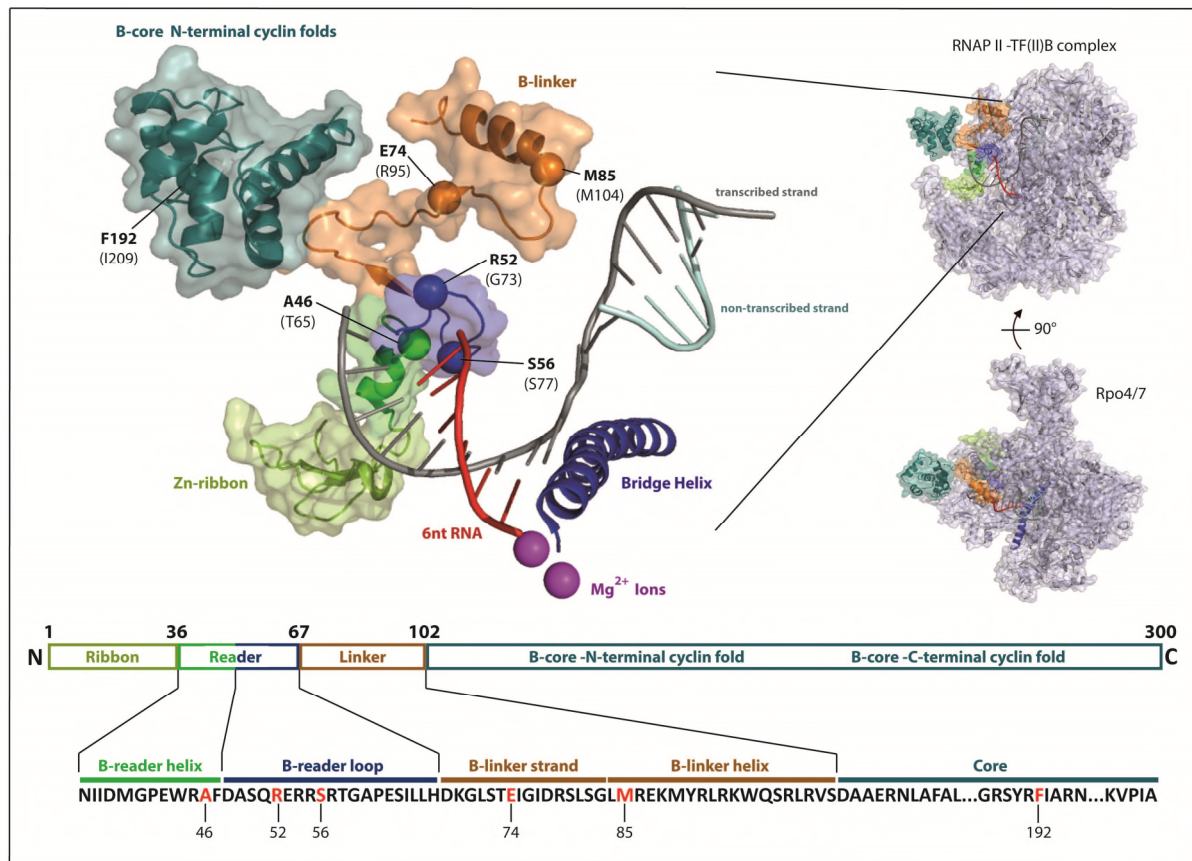
Figure 29: Localization of the selected TFB-Bpa variants in the initially transcribing complex (PDB: 4BBS). TFIIB domain organization is depicted in different colors: N-terminal cyclin fold of the B-core domain (cyan), B-linker (orange) and Zn-ribbon (lime green). The B-reader consists of the helix (green) and the loop (blue) and is located in proximity to the t-strand (grey). Bpa-substitutions are highlighted as dots, the position for *Pfu*TFB is given in bold letters, and corresponding *Sc*TFIIB positions are in brackets, respectively. The active site of the RNAP II is indicated by Mg^{2+} ions (magenta), the RNA (red) and the bridge helix (dark blue).

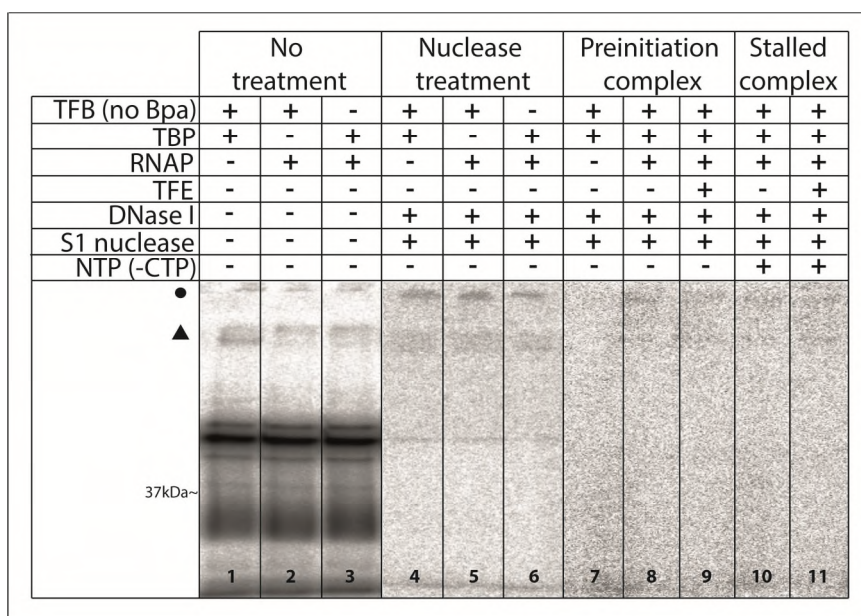## 2. Specificity of UV crosslinking experiments

The exposure of the TFB-Bpa proteins to UV-light generates a covalent crosslink to C-H bonds in proximity to the unnatural amino acid. Site-specifically radiolabeled DNA templates were used to detect specific interactions of TFB with DNA (Figure 23). To ensure the correct incorporation of the [α-^{32}P] dNTP at the respective site, small volumes of the samples were taken after the distinct labeling steps and loaded on a gel (Figure 30 C). Three signals were observed for each template on the gel, whereas the lower bands are from samples after the incorporation step, indicating that DNA is labeled. These bands of the respective gdh-C cassettes differ in the height because of the different template length and the corresponding oligonucleotides in this example. After addition of unlabeled dNTPs, bands occurred at the top of the gel, demonstrating that the template is completely double-stranded. After EcoRI treatment the signals were visualized lower on the gel which indicated that the template DNA was successfully released from magnetic particles. These DNA templates were used in crosslink reactions together with TFB-Bpa variants.
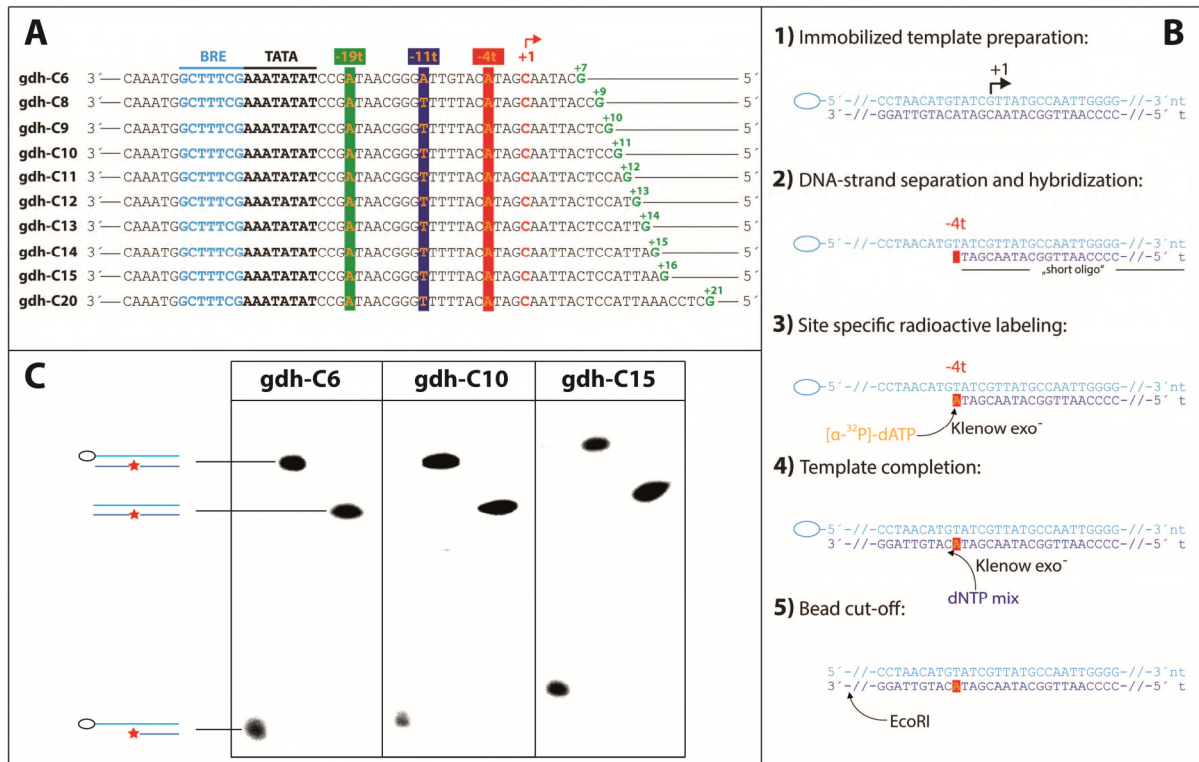
Figure 30: Preparation and overview of radiolabeled gdh-C cassettes. **A)** The transcribed strands of the particular cassettes are shown in 3´ to 5´ direction. All cassettes contained the strong glutamate-dehydrogenase (gdh) promoter of *P. furiosus*. The BRE sequence is depicted in light blue, the TATA box in black bold letters. The transcription start site is indicated with a red bold letter (+1 C), whereas the first guanine is in bold dark green. Sites for the incorporation of the radiolabeled nucleotide are highlighted in red (-4t), dark blue (-11t) and green (-19t). **B)** Schematic representation of the incorporation of a radiolabeled nucleotide to a specific site on the DNA. Templates were prepared by PCR and 5´-biotinylated M13 primer to immobilize the template at the nt-strand (1). DNA strands were separated and hybridized with an oligonucleotide complementary to the nt-strand (2). Radiolabeled dNTP together with Klenow exo- were added to incorporate the nucleotide site-specifically at the DNA t-strand (3). The template was completed by adding a full set of dNTPs (4), and at least the magnetic beads were removed by EcoRI digestion. **C)** Specificity of the radiolabeled nucleotide incorporation was analyzed on a 7M urea 12% PA gel. Semi-hybridized and labeled fragments (lower signals), strand-completed DNA (upper signals) and EcoRI digested fragments (middle signals) are shown for the cassettes gdh-C6, gdh-C10 and gdh-C15 as examples.

To identify a specific interaction between TFB and the radiolabeled site of the DNA, samples were treated with different enzymes and detergents. Dnase I and S1 nuclease were used to digest DNA and RNA, and the treatment with SDS denatured all proteins and protein-protein interactions of the complex. The resulting TFB-Bpa protein bound to a small fragment of radiolabeled DNA was analyzed using SDS-PAGE, whereas the TFB-DNA complex appeared as a distinct band at the height of 37kDa after the autoradiogram, which is the size of *Pfu*TFB. To verify that unspecific interactions were correctly digested, an experiment was performed with several samples containing different sets of proteins (Figure 31) together with a gdh-C6 cassette containing a [α^{32}P] dATP four nucleotides upstream the transcription start site (-4t). Control samples which lacked RNAP, TBP or wtTFB were exposed to UV light at a wavelength of 300nm and separated using SDS-PAGE without treatment (Figure 31 lanes 1 - 3). Several bands appeared on the gel, indicating unspecific interactions of the proteins to radiolabeled DNA. The same samples were treated with the enzymes Dnase I, S1 nuclease,

and the detergent SDS after UV light exposure (lane 4 - 6). No signals were detected on the gel at the respective sites, demonstrating the elimination of all these unspecific interactions. To show that wtTFB, which does not contain a Bpa photo crosslinker, does not bind to DNA nonspecifically in the preinitiation complex, samples containing respective transcription factors were exposed to UV light (lanes 7 - 9) and analyzed by SDS-PAGE. The results demonstrated that wtTFB does not interact to DNA in the DNA/TBP/TFB complex, as well as

	No treatment			Nuclease treatment			Preinitiation complex			Stalled complex	
TFB (no Bpa)	+	+	-	+	+	-	+	+	+	+	+
TBP	+	-	+	+	-	+	+	+	+	+	+
RNAP	-	+	+	-	+	+	-	+	+	+	+
TFE	-	-	-	-	-	-	-	-	+	-	+
DNase I	-	-	-	+	+	+	+	+	+	+	+
S1 nuclease	-	-	-	+	+	+	+	+	+	+	+
NTP (-CTP)	-	-	-	-	-	-	-	-	-	+	+

Figure 31: Digestion efficiency of the crosslink reactions with gdh-C6 cassette radiolabeled at -4t. All samples were separated using 12% SDS-PAGE and analyzed by autoradiography. The samples in the first three lanes contained different compositions of the transcription factors and RNA polymerase without nuclease treatment (No treatment) to identify unspecific interactions. The same samples were used in the lanes 4-6 but Dnase I, S1 nuclease and SDS were used (Nuclease treatment). Samples containing TBP/wtTFB/DNA-complexes (lane 7), and preinitiation complexes in absence (lane 8) and presence of TFE (lane 9) as well as samples of stalled complexes at position +6 with and without TFE (lane 10, 11) do not show a signal on the gel except weak signals marked with a triangle and a circle. These signals occurred in a factor-nonspecific manner, indicating background.

in preinitiation complexes with and without TFE. Also no signal was observed in stalled complexes at register +6 in absence and presence of TFE (lane 10, 11), indicating that wtTFB does not interact nonspecifically with the radiolabeled DNA. These results demonstrated that both the specific incorporation of radiolabeled nucleotides into DNA and the elimination of unspecific signals in open preinitiation complexes and stalled complexes after crosslink reactions worked specific. With this experimental setup crosslinking experiments were performed with the selected TFB-Bpa mutants.

### 3.   Crosslinking experiments in the preinitiation complex

It was shown that the eukaryotic TFIIB B-reader domain is in proximity to the t-strand and to the active center of the polymerase in crystal structures (Bushnell et al. 2004; Sainsbury et al. 2013). Therefore crosslinking experiments were performed with gdh-C6 cassettes containing a radioactive nucleotide at position -4 relative to the transcription start site on the transcribing strand. For each Bpa position three samples were prepared, whereas the first reaction lacked the RNAP to mimic a DNA/TBP/TFB-Bpa complex (Figure 32). In this ternary complex DNA is in a closed conformation and therefore the radiolabeled site of the DNA is out of distance to the Bpa. This reaction was performed as a negative control to verify the

specificity of the crosslink reactions. The second sample contained RNAP to form open preinitiation complexes in which the photo crosslinker should be in proximity to the labeled site on the t-strand. The third samples contained additional TFE which is usually part of the PIC *in vivo*. The results showed that the wtTFB does not crosslink to the labeled DNA nonspecifically, as it was expected (Figure 32 A). In contrast, in reactions containing the B-reader variants A46Bpa, R52Bpa and S56Bpa a signal appeared on the gel at the size for in TFB at approximately 37kDa, indicating the TFB-DNA complex. This signal was only present in reactions containing RNAP, suggesting that the crosslink of the B-reader variants is specific for open complexes. The signal intensity of the crosslinked TFB-DNA complexes increased with addition of TFE, which indicates that more complexes can be crosslinked in presence of TFE, or addition of TFE reduces the distance between the Bpa and the label, possibly due to a better stabilization of the transcription bubble within the complex. The B-reader loop mutation R57Bpa was also used in the reactions as a negative control, because this amino acid substitution showed massive deficiencies in promoter opening and transcription in the analysis. No signal was detected on the gel for this TFB position, indicating that open complex formation failed. In addition to this crosslink, the B-linker positions E74Bpa and M85Bpa were also used in reactions containing a gdh-C6 template with a label at -4t. The results showed that no contact between the Bpa positions and DNA can be detected. The B-core position F192Bpa showed slight background in the crosslinks but the fact that the signal decreased with addition of RNAP and TFE, as well as the large distance of the corresponding amino acid to the labeled site on the DNA in the crystal structure, suggested that the signal is highly nonspecific.

In a modelled open complex structure of the yeast PIC the TFIIB B-linker strand is in close proximity to the t-strand approximately 11 nucleotides upstream the transcription start site (Kostrewa et al. 2009) (Figure 32 D). To test if the t-strand is located in proximity to the archaeal TFB B-linker strand a gdh-C6 DNA template was designed containing a radiolabeled nucleotide at position -11t. Crosslink experiments in presence of TFE were performed with the selected TFB B-reader variants A46Bpa, R52Bpa and S56Bpa together with the TFB B-linker mutations E74Bpa and M85Bpa as well as wtTFB as a negative control (Figure 32 B). The B-reader mutations do not crosslink to DNA radiolabeled at -11t, indicating high specificity of the crosslinking system. A strong crosslink signal for E74Bpa, but not for M85Bpa at the expected size for TFB was observed in this experiment, indicating a close location of E74 to the t-strand in the open complex. To test the specificity of the contact between E74Bpa and DNA a crosslink experiment was performed in absence of RNAP, in absence of TFE and in presence of all factors (Figure 32 B). The result of this assay clearly demonstrated that E74 crosslinked specifically to DNA only in the open complex and the signal is increased by addition of TFE. This finding suggests that the upstream end of the t-strand transcription bubble is located next to the B-linker strand, whereas the B-linker helix position M85Bpa seems to be out of distance to crosslink DNA at this position. The fact, that E74Bpa crosslinks to DNA labeled at -11t but not to DNA labeled at -4t additionally emphasize the specificity of the crosslink system used in this study. Because of the close location of TFB-E74 to the very upstream part of the transcription bubble it was interesting to see if the B-linker mutations are able to crosslink also to the non-transcribed DNA strand. Therefore, crosslinking experiments were performed using a gdh-C10-8nt template, which contains a label at position -8 on the nt-strand relative to the transcription start site, together with the B-linker positions E74Bpa and M85Bpa, as well as the controls TFB-F192Bpa of the B-core, TFB-R52Bpa of the B-reader and wtTFB (Figure 32C).
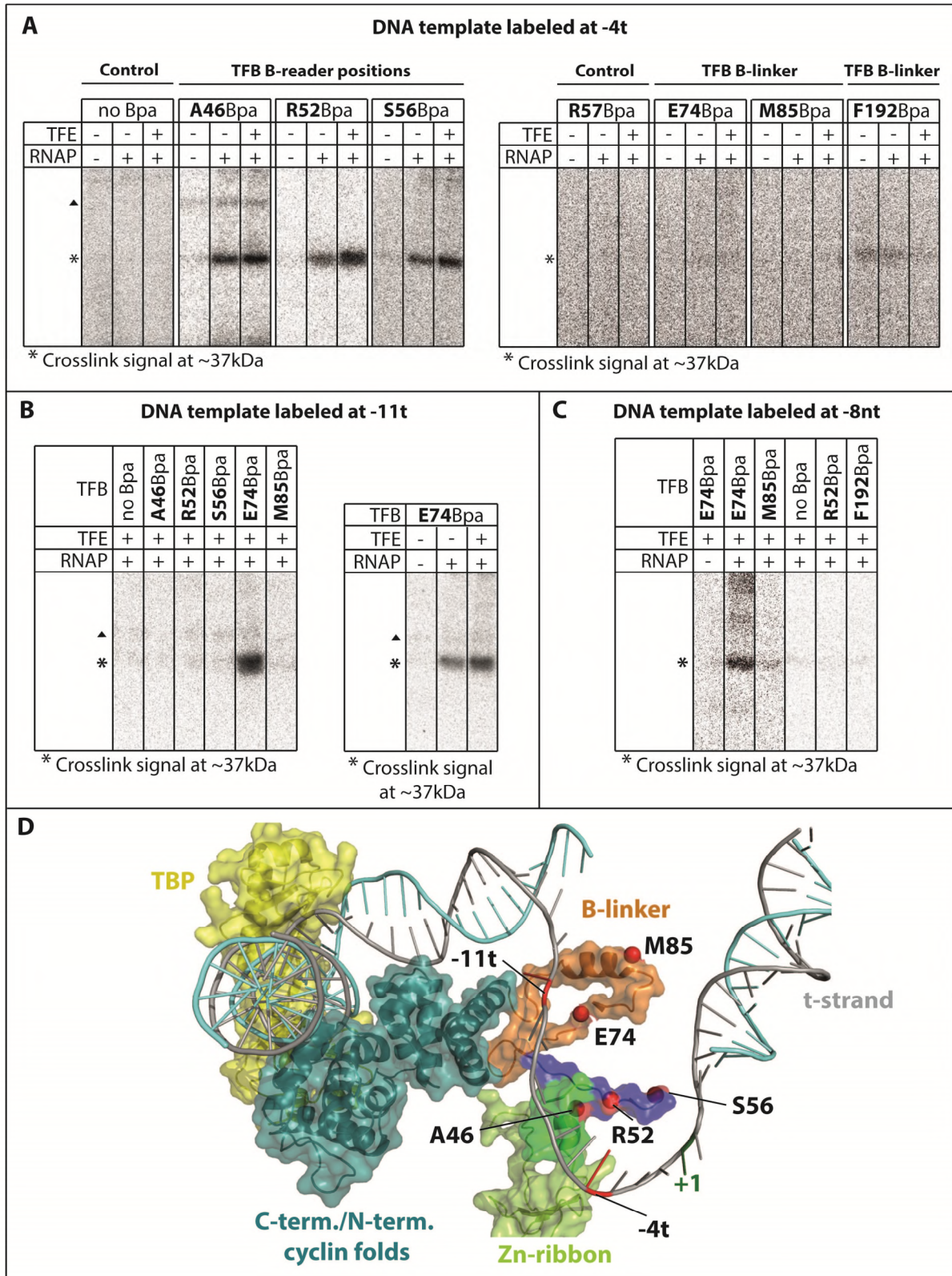
Figure 32: TFB-DNA crosslink reactions in preinitiation complexes. **A)** The B-reader variants A46Bpa, R52Bpa and S56Bpa as well as the linker mutations E74Bpa and M85Bpa, and the core mutation F192Bpa were used in crosslink reactions together with a gdh-C6 DNA template labeled at -4 t-strand. TFB without Bpa and R57Bpa were used as negative controls. The TBP/TFB/DNA-complexes for each mutation are represented in the first lanes and the open complexes in the second lanes, respectively. TFE was additionally present in the samples shown in the third lanes. A radioactive signal was observed on the 12% SDS-PAGE at a size of 37kDa for the B-reader mutations only in presence of

RNAP, which is the result of a covalent bond between TFB and the radiolabeled DNA. Signals marked with a triangle appear in some cases but not factor specific, indicating an unspecific signal. **B)** Three B-reader mutations A46Bpa, R52Bpa and S56Bpa together with wtTFB and the two TFB B-linker mutations E74Bpa and M85Bpa were used in crosslink experiments with gdh-C6 labeled at -11t. TFE and RNAP are present in every sample. Only E74Bpa showed a specific crosslink signal at the expected size of 37kDa on the SDS-PAGE. Signals marked with a triangle are unspecific. TFB-E74Bpa crosslinked to DNA labeled at -11t in presence of RNAP (middle lane) and with additional TFE (last lane), but not in the TBP/TFB/DNA-complex, indicating a specific interaction of E74Bpa to DNA. **C)** The B-linker mutations E74Bpa without RNAP (first lane) and with RNAP (second lane), together with M85Bpa, as well as the wtTFB, the B-reader mutation R52Bpa and the B-core mutation F192Bpa were crosslinked to DNA labeled at -8nt. Only E74Bpa showed a signal at the expected height for TFB, and only if RNAP is present, indicating a specific interaction of E74 and the nt-strand in open complexes. **D)** Postulated eukaryotic yeast open complex model containing t-strand region -11t (PDB:3K1F, (Kostrewa et al. 2009)). TBP is colored in yellow, and TFB domains are depicted in different colors (B-core in cyan; B-reader helix in green, B-reader loop in blue, B-linker in brown) whereas RNAP is excluded. TFB-Bpa positions are shown as red dots, whereas the radiolabeled sites are highlighted in red.

The results showed that TFB-E74Bpa crosslinked to DNA labeled at this specific site only in presence of RNAP, indicating a specific interaction with DNA in the open complex only. In contrast, the other positions tested showed no formation of a covalent crosslink between the Bpa position and the radiolabeled DNA, demonstrating that E74Bpa is the only amino acid position of the selected variants able to crosslink to the nt-strand. Therefore the location of E74 is likely between the separated strands at the upstream edge of the transcription bubble.
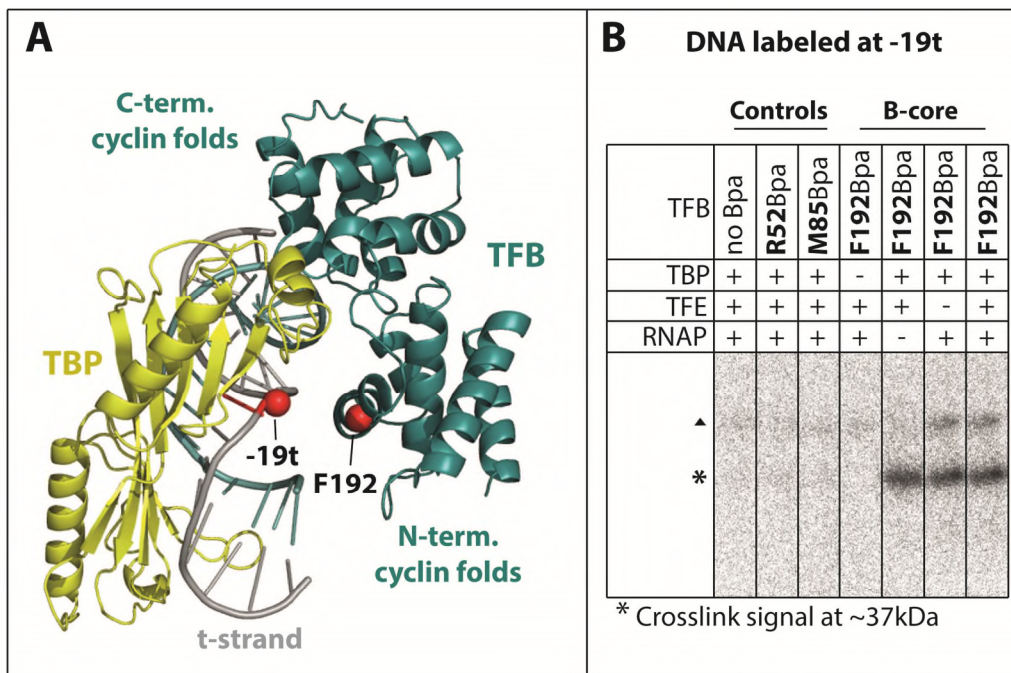
Figure 33: TFB B-core mutation F192Bpa crosslinks to DNA radiolabeled at -19 t-strand. **A)** Crystal structure containing TATA-box, TFB B-core (cyan) and TBP (yellow) of *P. woesei* (PDB: 3AIS) (Kosa et al. 1997). Amino acid F192 and DNA at position -19t are indicated by red dots. **B)** TFB without Bpa, TFB-R52Bpa, TFB-M85Bpa and TFB-F192Bpa were used in crosslink reactions with a gdh-C6 cassette containing a radiolabeled nucleotide 19 bases upstream of the transcription start site. TFB-F192Bpa showed the expected radioactive signal on the SDS-PAGE at a size of 37kDa in the TBP/TFB/DNA-complex and in the open complex with and without TFE, whereas the other mutations and wtTFB do not crosslink to DNA at this site, indicating a specific interaction of F192 and DNA. Signals marked with a triangle indicate unspecific background.

Position F192 of the TFB B-core domain was shown to be in proximity to the DNA in crystal structures of the related *P. woesei* ternary TATA-box/TBP/TFB B-core complex (Kosa et al. 1997) and also the corresponding amino acid I209 of the yeast TFIIB in modelled open and closed complexes (Kostrewa et al. 2009). Therefore crosslink experiments were performed with the TFB-F192Bpa mutation and a DNA template containing a radiolabeled nucleotide 19 base pairs upstream the transcription start site at the transcribing strand. The results of the experiments showed that wtTFB, the B-reader position R52Bpa and the B-linker position M85Bpa do not interact with DNA labeled at this site (Figure 33 B). The B-core mutation F192Bpa does not contact DNA in absence of TBP, but in DNA/TBP/TFB complexes, as well as in preinitiation complexes with RNAP and/or additional TFE, indicating a specific interaction between F192 and DNA at -19t. It is also interesting to note that addition of TFE to the open complex does not lead to an increased signal as it was shown for the B-reader variants (Figure 32 A). This finding indicates that TFE does not lead to more PICs due to the stabilization effect of the transcription bubble in the open complex. Moreover, TFE induces structural changes at the active site which leads to stronger crosslinks for the B-reader mutations.

Taken together, mapping of amino acids of TFB to respective sites on the DNA revealed that the location of selected TFB variants is nearly similar to the location of corresponding amino acids of eukaryotic TFIIB in crystal structures. This finding suggests a related topology of the transcription factor B within the open preinitiation complex in the archaeal transcription system in comparison to the eukaryotic transcription system.

4. Crosslinking experiments in stalled transcription complexes

To analyze possible structural rearrangements of the TFB B-reader domain during transition from transcription initiation to early elongation crosslinking experiments were performed using TFB B-reader mutations A46Bpa, R52Bpa and S56Bpa on stalled transcription complexes. The gdh-C cassettes allowed pausing of transcribing complexes, whereas cassettes gdh-C6, gdh-C8, gdh-C9, gdh-C10 and gdh-C15 were used in the crosslink experiments to stall complexes at registers +6, +8, +9, +10 and +15 (Figure 34 A). Each template was individually radiolabeled at position -4t. Signal intensities of the resulting TFB-DNA contacts in stalled complexes were compared with those derived from preinitiation complexes. Altered signal intensities of the crosslinks between PICs and stalled complexes can be interpreted as a change in the distance between Bpa in the protein and the labeled DNA, indicating structural changes of TFB and/or DNA.

To validate the correct stalling of transcription complexes, transcription assays with radiolabeled UTP without CTP were performed to see if the complexes for each TFB mutant can be positioned correctly on each template. First, the incubation temperature was adjusted to achieve best stalling results of the complexes to avoid falsified results (Figure 34 B). Several combinations of the pre-incubation and stalling temperatures were tested using wtTFB, whereas the samples were analyzed on a high percentage gel to validate specificity of stalling. The results showed that stalling works optimally at an incubation temperature of 80°C, because additional bands at the upper part of the gel did not appear. In contrast, other temperatures tested showed signals at the top site of the gel, indicating that the RNA polymerase read over the -C position of the template, leading to a minor population of run-off transcripts. Using these conditions a chase-experiment was performed to analyze if the stalled complexes were transcriptionally active (Figure 34 C). Reactions were split into two samples after stalling the complexes, whereas CTP and non-labeled UTP was added to the second reaction. More than 80-85% of the RNAs disappeared in samples were CTP and non-labeled UTP was added and a signal at the upper side of the gel was detected,
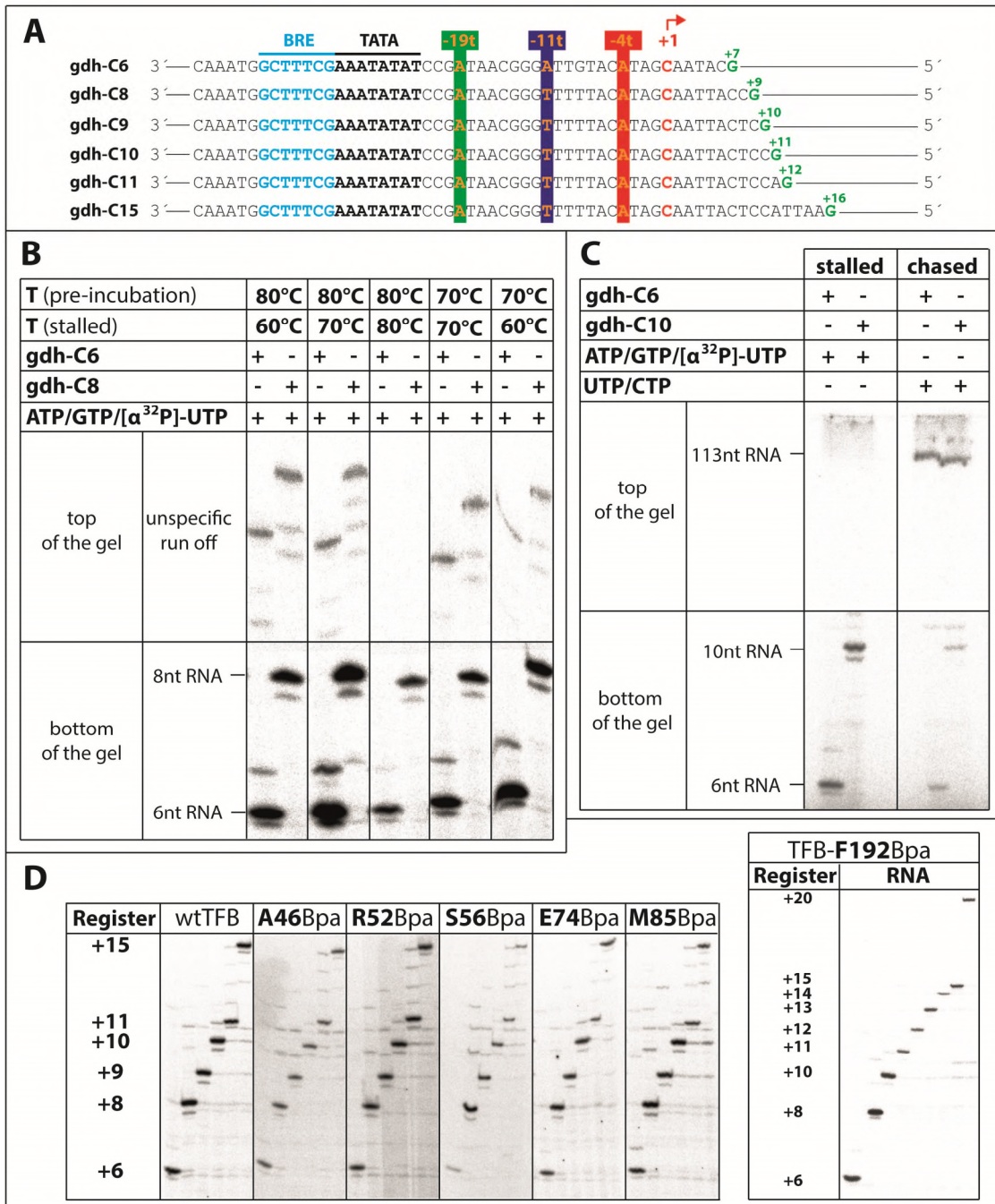
Figure 34: Analysis of stalled transcription complexes. **A)** Overview of the templates used for stalling transcription complexes. All templates contained a strong gdh promoter consisting of a BRE (light blue) and a TATA-box element (black). The t-strands are shown in 3′ to 5′ orientation which is the transcription direction. TSS is indicated by a red arrow at +1. Labeling sites for incorporation of radio labeled dNTPs are highlighted in green (-19t), blue (-11t) and red (-4t). Positions for stalling are shown as green bold letters. **B)** Conditions to stall complexes were optimized using different pre-incubation and stalling temperatures. Gdh-C6 and gdh-C8 templates were used in reactions lacking CTP. The combination of 80°C pre-incubation and 80°C for stalling showed the best results on the gel because unspecific signals at the top of the gel were not detected. **C)** Stalled transcription complexes at register +6 and +10 were chased by adding UTP and CTP in molar excess. Quantification of the signals showed that 80-85% of the 6-10nt RNAs can be extended to the full run-off transcript after 20 minutes incubation time, indicating that the stalled transcription complexes are transcriptionally competent. **D)** Selected TFB-Bpa linker and reader variants were stalled at registers +6, +8, +9, +10, +11 and +15, and the B-core F192Bpa was stalled at +6, +8, +10 to +15 and +20 to analyze if complexes initiated

with these TFB mutants are able to position complexes correctly at the respective registers. Main RNA signals were observed at the same heights like RNAs of the wtTFB, indicating that all positions tested are able to stall the polymerase correctly.
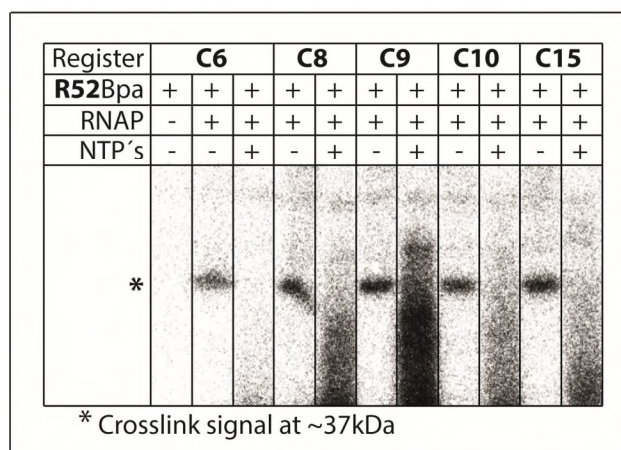
indicating formation of run-off products. The result showed that the stalled transcription complexes are transcriptionally competent after 20 minutes incubation time, which is the time of the UV exposure in the crosslinking experiment. In a last experiment the TFB-Bpa mutations were used in stalling experiments at registers +6, +8, +9, +10, +11 and +15 (Figure 34 D). Register +11 was used as an additional control. All variants tested showed a distinct RNA pattern on the gel comparable to the wtTFB, indicating that stalling with the TFB mutants is applicable.

Based on this stalling procedure, the TFB B-reader variants were used in crosslinking reactions on preinitiation and stalled complexes to identify possible structural rearrangements of the B-reader domain (Figure 35). The crosslink reaction on ternary DNA/TBP/TFB-Bpa complexes was performed for every TFB position tested at the gdh-C6 template and is shown on the left lanes of each gel (Figure 35 A, B, C). No signal was detected in these experiments, indicating a specific crosslink reaction to the t-strand in open complexes. At registers +6, +8, +9, +10 and +15 crosslinking experiments were performed with TFE in absence and presence of NTPs without CTP, leading to the formation of preinitiation and stalled complexes. Signals derived from crosslink reactions on stalled complexes were compared with signals of preinitiation complexes of the respective cassette. Altered signal intensities indicating changes in the distance between Bpa and the radiolabeled DNA, which allows conclusion about structural rearrangements of the B-reader domain. At register +6 no change in the signal intensity of the stalled complexes in comparison to the preinitiation complexes was detected for A46Bpa and R52Bpa, whereas S56Bpa showed a slightly reduced signal of 85%. At register +8 R52Bpa and S56Bpa showed unaltered patterns in comparison to register +6, whereas the signal of TFB-A46Bpa decreased to 72%, indicating an increased distance of the B-reader helix to DNA (Figure 35 A). At register +9 A46Bpa and S56Bpa had stronger signals in comparison to register +8. In contrast, the signal intensity of R52Bpa started to decrease at this point. The signal intensity of all selected TFB positions are markedly reduced at register +10 and remained unchanged low at register +15. This finding suggests a translocation of the TFB B-reader domain, which is likely a displacement of the B-reader domain. Signals are still present on the gels, suggesting that the residual signal derived from stalled complexes at register +10 is background. It is possible that not all preinitiation complexes were stalled and are still at the promoter start site, or the B-reader domain is displaced but is still in a suitable distance for a weak crosslinking reaction. To examine if the TFB B-reader domain also showed signals in crosslinking experiments performed under run-off conditions, R52Bpa was used in reactions containing the respective gdh template and a full set of NTPs. No signal was observed at the height of TFB on the gel, indicating that the B-reader variant does not crosslink to DNA under run-off conditions.

To investigate if the observed reduction in signal strength is based on structural rearrangements of the B-reader domain and not caused by a release of TFB from the complex, crosslinking experiments were performed with the B-core variant F192Bpa. The use of this TFB position has the advantage that rearrangements happening at the active site of the RNA polymerase should not affect the TFB-F192Bpa-DNA interaction. Therefore changes in resulting signal intensities between stalled and preinitiation complexes indicate an altered number of TFB proteins present at the DNA.
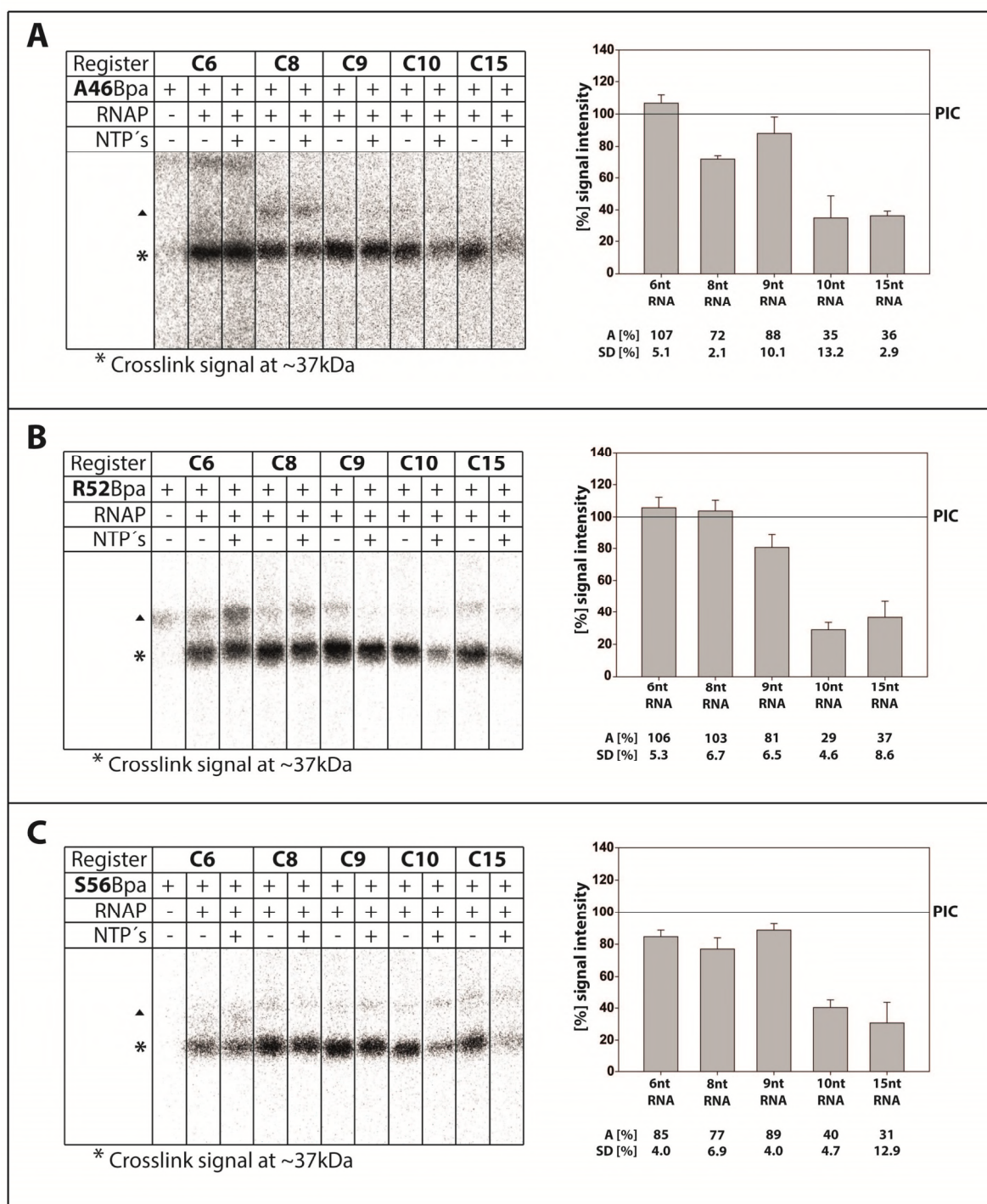
Figure 35: Crosslinking experiments with the TFB B-reader mutations A46Bpa, R52Bpa and S56Bpa and -4t radiolabeled DNA in stalled complexes halted at +6, +8, +9, +10 and +15. **A)** Crosslink reactions of the B-reader helix mutation A46Bpa. **B)** Crosslink reactions of the B-reader loop mutation R52Bpa. **C)** Crosslink reactions of the B-reader loop mutation S56Bpa. For each figure the TBP/TFB/DNA complex is shown in the left lane of the results derived from cassette gdh-C6 as a negative control. The second lane represents the preinitiation complex, the third lane the stalled complex. For the results with cassettes C6 to C15 the signal derived from preinitiation complexes (left lanes) and the signals derived from respective stalled complexes are shown. Signals marked with a triangle derived from factor unspecific interactions and are not present in every lane. Three individual experiments were performed, and the signal intensity of stalled complexes were quantified and compared to the respective signal intensity of preinitiation complexes. The standard deviation (SD) and the average (A) were calculated and are summarized in a bar diagram.

Register	C6			C8		C9		C10		C15	
R52 Bpa	+	+	+	+	+	+	+	+	+	+	+
RNAP	-	+	+	+	+	+	+	+	+	+	+
NTP´s	-	-	+	-	+	-	+	-	+	-	+

* Crosslink signal at ~37kDa

Figure 36: TFB-R52Bpa does not crosslink to -4t radiolabeled DNA under run-off conditions. TFB-R52Bpa crosslink specifically to DNA labeled at -4 t-strand in open complexes but not in the TBP/TFB/DNA-complex. Run-off conditions were generated using a complete set of NTPs, whereas TFB-R52Bpa does not crosslink to DNA, as no signal was observed for the respective cassette at the expected size of 37kDa. Heparin was used as competitor to prevent re-initiation of TFB to the promoter site.

To investigate the presence of TFB at register +6, +8, +10 and +15, complexes were stalled at the respective sites with individually labeled gdh-C templates at position -19t. The results showed that in three individually performed crosslinking experiments the signal intensity of stalled complexes do not differ from signal intensities of preinitiation complexes, indicating that the number of TFBs does not change at the respective registers (Figure 37). Based on this finding, it can be concluded that the markedly reduced signal intensities observed for the TFB-Bpa variants of the TFB B-reader domain with DNA labeled at -4t at register +10 (Figure 35) derived from specific transitions of the reader domain. This transition is likely due to a displacement of the B-reader from DNA, which results in collapse of the transcription bubble. The use of this TFB-F192Bpa mutation further allows analyzing the position on the DNA, where TFB release takes place. As it was predicted from the initially transcribing complex structure, RNA clashes with the Zn-ribbon of TFIIB at a length of 12 - 13nt resulting in release of the transcription factor. To investigate if archaeal TFB dissociates from the complex at this position, gdh-C11 to gdh-C14 templates were designed. Stalling of TFB-F192Bpa was analyzed using [α-32P] UTP (Figure 37 D), whereas a RNA ladder was observed, indicating that complexes can be positioned at the respective sites. Crosslinking experiments using TFB-F192Bpa and the generated templates containing a radioactive label at -19t were performed, and in addition to the usually used competitor heparin, poly-dAdT was used to trap TFB to prevent re-association of TFB to the promoter site after a possible release event. The results showed that the signal intensities of stalled complexes do not differ from signal intensities of preinitiation complexes at registers +11, +12, +13, and +14. In contrast, TFB tended to be released from register +15 onwards, is reduced at +20 and completely absent in crosslinking experiments under run-off conditions. The results suggest that TFB indeed is released from the complex, as the signal intensity started to decrease from register +15, but it seems that this is not an instantaneous event, as it was expected.
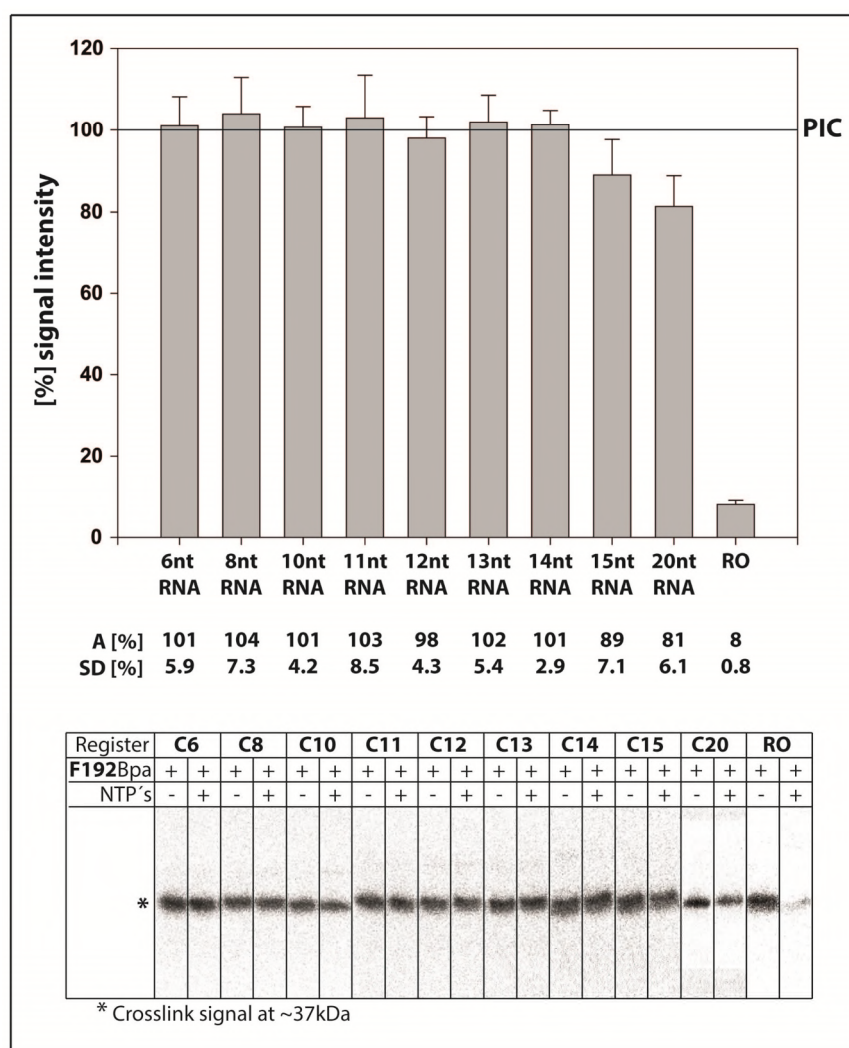
Figure 37: TFB is present at register +6 to +14, started to decrease at register +15 to +20, and is absent under run-off conditions. TFB-F192Bpa was crosslinked in preinitiation complexes (left lanes for each cassette) and in stalled complexes at the respective site (right lane for each cassette) with gdh-C cassettes containing a radiolabeled nucleotide at position -19 t-strand. The crosslink under run-off conditions was performed on the gdh-C20 cassette. All reactions were performed in three individually experiments, whereas the signals derived from crosslink experiments in halted complexes were compared to the signals of the respective PIC. The standard deviation (SD) as well as the average (A) was calculated and summarized in the bar diagram.

5.   Summary of the crosslinking experiments

To analyze *Pfu*TFB and its interactions to DNA during transcription initiation and transition to early elongation, TFB-Bpa variants of the B-reader domain (G41 to R57), the B-linker domain (E74 and M85) and the B-core domain (F192) were successfully created. The mutants were screened using *in vitro* transcription assays to validate the applicability of altered TFB proteins in crosslinking experiments. Based on the position within TFB and the overall quality in transcription six positions were selected: three amino acids of the B-reader domain (TFB-A46Bpa of the reader helix, TFB-R52Bpa and TFB-S56Bpa of the reader loop), two of the B-linker (TFB-E74Bpa and TFB-M85Bpa), and the B-core variant TFB-F192Bpa (Figure 28). The selected TFB-Bpa variants were able to form preinitiation complexes (Figure 24), open complexes (Figure 27), and RNA polymerases initiated with these TFB proteins were able to form the first phosphodiester bond (Figure 25) and run-off transcripts (Figure 26) in the

respective experiments. In preliminary experiments stalling of transcribing complexes was shown to work optimally at 80°C, whereas complexes were transcriptionally competent after 20 minutes incubation time (Figure 34). Stalling of the TFB-Bpa variants on respective gdh-C templates was shown to be specific, whereas the successful site-specific radioactive labeling of the DNA templates was verified (Figure 30). The used crosslinking system worked specific, because nonspecific interactions between proteins and the labeled DNA were successfully eliminated after exposure to UV light (Figure 31). Crosslinking experiments in preinitiation complexes with DNA labeled at -4t showed that the B-reader positions TFB-A46Bpa, TFB-R52Bpa and TFB-S56Bpa can be covalently crosslinked to DNA only in open complexes when RNAP is present (Figure 32). Control assays with B-linker and B-core positions, as well as with wtTFB and TFB-R57Bpa, a position insufficient for promoter opening, showed no or unspecific background. In crosslink reactions with -11t labeled DNA and -8nt labeled DNA a specific contact between the linker TFB-E74Bpa and DNA was observed. The TFB-F192Bpa mutation crosslinked to -19t labeled DNA specifically in TBP/DNA/TFB-Bpa complexes, as well as in open complexes (Figure 33). Crosslinking reactions on stalled complexes at registers +6, +8, +9, +10 and +15 with DNA labeled at -4t revealed that the signal intensity of A46Bpa at register +8, and S56Bpa at registers +6 and +8 were reduced, indicating interactions with nascent RNA (Figure 35). The signal intensity at register +10 was markedly reduced for all TFB B-reader variants, suggesting a displacement of the reader domain, resulting in collapse of the transcription bubble. The presence of TFB at complexes was verified using crosslinking experiments with TFB-F192Bpa and DNA labeled at -19t. It was shown that the number of TFBs did not differ at registers +6 to +14, demonstrating that the observed decreased signal intensities at register +10 are specific for the TFB B-reader domain. The crosslinking experiments further revealed that the number of TFBs tended to decrease from register +15 onwards, is reduced at +20, and absent in reactions under run-off conditions, indicating that interactions of TFB within the complex were destabilized at register +14, resulting in a TFB release.

V. Discussion

A. A possible role for RPA during transcription elongation

*Pfu*RPA was used in different *in vitro* experiments to analyze its possible function during the transcription process. The protein showed the expected preference to ssDNA (Figure 9 A), but no effect was observed in abortive transcription assays (Figure 9 B). In addition, western blot experiments and EMSAs on initiation complexes were performed (data not shown), but no evidence was found that RPA is part of initiating complexes. In contrast, *Pfu*RPA showed activity in chase experiments of stalled complexes using a 4kb plasmid in which the formation of intermediates was prevented (Figure 10 B). *Pfu*RPA further increased transcription processivity of stalled complexes in time-dependent transcription reactions (Figure 11). In accordance to the presented results of Pluchon et al., *Pfu*RPA showed formation of transcripts with 2.9-fold increase, indicating that *Pfu*RPA functions in the same manner like *Pab*RPA (Pluchon et al. 2013). Therefore *Pfu*RPA likely functions during elongation of transcription and interacts with RNAP in a stabilizing manner, which further increases its processivity.

Richard et al. reported a stimulatory effect of the *S. solfataricus* SSB under TBP limiting conditions (Richard et al. 2004). In these reactions TBP was reduced to 20% of the regular concentration and they observed that transcription takes place only in presence of RPA. This finding would suggest that RPA functions in RNAP recruitment, which leads to more initiation events and therefore increase transcription output. The idea, that RPA recruits the RNA polymerase is supported by the finding that RPA is associated with RNAP in solution, which was shown by (Komori, Ishino 2001). *Pfu*RPA was also tested under TBP-limiting conditions and showed similar effects as described in (Richard et al. 2004) (data not shown). Because there is no evidence for a participation of *Pfu*RPA at the preinitiation complex, but transcription is also stimulated under TBP limiting conditions, an increased processivity of the RNAP is more likely than a role in RNAP recruitment, which would explain the higher number of formed transcripts in these assays. However, based on the performed experiments a role in RNAP recruitment can not be excluded.

Moreover, the herein presented results suggest a stimulation of *Pfu*RPA on transcription, likely due to a stabilization effect and an increased processivity of the RNA polymerase. Increased transcription speed as well as the prevention of pausing events of the RNA polymerase can explain the formation of more transcripts during transcription. However, RPA was shown to bind ssDNA, and if RPA functions during elongation, this factor should interact with the single stranded transcription bubble during elongation, as it was hypothesized by (Sikorski et al. 2011). It was tried to show the presence of RPA in the elongation complex using western blot experiments as well as EMSAs with radio- and fluorescently labeled DNA templates, and stalled complexes with radiolabeled RNA (data not shown), but none of these experiments could verify the presence of RPA in the elongation complex, possibly due to a weak interaction to the RNAP in the used *in vitro* assays.

However, to deepen the understanding of the molecular mechanism of RPA during transcription, $KMnO_4$ footprints on preinitiation complexes as well as on stalled elongation complexes might elucidate a direct interaction with the single-stranded region of the transcription bubble. This assay should clarify the proposed DNA melting effect of the SSB during initiation, which explain the increased transcription rates in the Sulfolobus system (Richard et al. 2004), and additionally should show if RPA can increase the stability of single-stranded regions during elongation of transcription due to a direct interaction with the transcribed or non-transcribed strand in transcribing complexes, as it was postulated for the

eukaryotic yeast system (Sikorski et al. 2011). Moreover, termination assays might also be helpful to investigate the function of RPA in the last stage of transcription. If RPA would support termination, RNAPs are more efficiently dissociated from DNA, which would increase recycling events to restart the transcription faster.

B.   Bending of DNA depends on the presence of TFB in *P. furiosus*

Bending of DNA at the promoter site is a prerequisite for the correct assembly of the preinitiation complex and therefore important for specific interactions between the transcription machinery and DNA (Nikolov et al. 1995). Bending results in a DNA conformation with a kink of approximately 90° angle, whereas general transcription factors are required to enable this conformational change (Juo et al. 1996). TBP, which recognizes the TATA element of the core promoter (Kim et al. 1993a), as well as TFB, which contacts TBP and DNA on both sides of the TATA (Renfrow et al. 2004), are necessary for this step. Gietl et al. demonstrated that different organisms follow different DNA bending pathways using single molecule analysis. DNA bending in the euryarchaeal organism *M. jannaschii* requires only TBP to bend DNA and addition of TFB does not change the observed DNA bending pattern. In contrast, in the crenarchaeal organism *S. acidocaldarius* presence of TBP does not bend DNA, but addition of TFB led to a high population of bent DNA and only a small group of non-bent DNA, indicating that DNA is in a kinked conformation only in presence of both factors. In addition, a three-step binding scenario was proposed for the eukaryotic organism *S. cerevisiae*. Here, addition of TBP results in two interconvertible high FRET states due to a step-wise binding mechanism of TBP to the promoter, and addition of TFIIB leads to fully bent DNA conformations (Gietl et al. 2014).

In order to investigate DNA bending behavior in the euryarchaeal *P. furiosus* transcription system, FRET measurements were performed using confocal and TIRF microscopy. To enable comparison with published results, the same SSVT6 promoter was used in this study as described in (Gietl et al. 2014). The results shown in chapter IV. B demonstrated that bending of DNA requires the action of TFB, as high FRET populations with a mean FRET efficiency of 40.7%±17.2% in confocal microscopy measurements, and 48.8%± 0.2% in TIRF microscopy measurements occurred only in presence of this transcription factor. The observed results show that DNA bending mechanism is not uniform within the euryarchaeal phylum, because the mechanism of *P. furiosus* seems to be more Crenarchaeota-like. However, binding to and bending of DNA relies on specific interactions of respective amino acids of TBP and TFB with the nucleic acids and therefore differences in the amino acid composition at the DNA-protein interfaces can influence the behavior of binding and bending of the DNA. Therefore complete preinitiation complex structures of members of the different phyla might be useful to understand the observed differences in bending behavior, but these structures are lacking. Another point to consider is the template which indeed contains TATA and BRE, but is a viral promoter (SSVT6). In addition, the measurements were performed at room temperature, and not at elevated temperatures. Therefore it remains speculative if the *P. furiosus* transcription would also show a TFB dependent DNA bending at physiological temperatures at specific *Pfu*-promoters or if TBP alone may enable bending. Nevertheless, the immobilized complexes containing bent DNA were stable in the TIRF measurements for >30 seconds until the donor dye bleached, and the efficiency value was very stable, showing that the formed complexes were highly stable and no dynamic switch or intermediate states were observed. This finding further indicates that TFB interact with the TBP-DNA complex in a stabilizing manner. It is also not clear if TBP and TFB interact with DNA in a stepwise manner as it was shown for the eukaryotic system (Masters et al. 2003; He et al. 2013), or if

both factors interact simultaneously with the promoter DNA. However, based on the relation to the eukaryotic transcription system, a stepwise binding in which TFB stabilizes the TBP-DNA interaction seems plausible, whereas bending of DNA depends on the presence of TFB in the euryarchaeal *P. furiosus* transcription system.

C. The charge distribution of the B-reader loop is important for the function of TFB

Since the first RNAP II - TFIIB co-crystal was resolved in 2004, it was shown that the TFIIB B-reader domain is located in proximity to the active site of the RNA polymerase II (Bushnell et al. 2004). The reader domain consist of the B-reader helix and the B-reader loop (Kostrewa et al. 2009), whereas it was shown that the helix is involved in TSS selection (Pinto et al. 1992; Li et al. 1994; Pardee et al. 1998; Kostrewa et al. 2009). Possible functions of the B-reader loop derived from structures of a modelled open complex and of an initially transcribing complex of yeast Pol II system (Kostrewa et al. 2009; Sainsbury et al. 2013). The structures located the B-reader loop closely to the transcribing strand of the DNA and to the active site of the polymerase, indicating a role in stabilization of the transcription bubble. In addition, a further role in RNA-strand separation was postulated via a charge-dependent mechanism (Sainsbury et al. 2013). To analyze the function of the B-reader loop domain in the related transcription system of *P. furiosus*, an alanine screen of corresponding amino acids was performed using different combinations of TFB B-reader-A substitutions to stepwise reduce the overall charge of the loop tip region to investigate impacts on transcription (Figure 15 B).

The results presented in chapter IV. C showed that all TFB variants tested form a preinitiation complex, but only R52A and E53R54A showed moderate activity, R54A weak and R52E53A increased or wtTFB activity in abortive and run-off transcription assays. Promoter opening in absence of TFE was sufficient for three of the TFB mutations (R52A, R52E53A, and E53R54A), whereas addition of TFE can compensate defects of some of the proteins tested, indicating that the B-reader loop is involved in stabilization of the transcription bubble. TFE can usually compensate defects of TFB during transcription (Werner, Weinzierl 2005), but abortive and run-off transcription assays in presence of TFE showed that in case of E53A, R54R55A and the LoopA substitution no or very low transcription was observed. The findings indicate an important role of this B-reader loop region in transcription initiation which possibly relies on its distinct charge distribution. The investigated B-reader loop tip region contains one acidic glutamic acid at position 53, whereas the other positions comprise a basic arginine at position 52, 54 and 55 (Figure 16). Multiple sequence alignments of this region revealed that all TFBs of the aligned organisms contain at least one negative charged amino acid at the corresponding E53 position, except TFB of *M. jannaschii*. Alanine screening of the *M. jannaschii* TFB of this domain revealed that substitution of the corresponding amino acids K87, I88, K89 and R90 (R52, E53, R54 and R55 of *Pfu*TFB) with alanine has no influence on abortive transcription assays except R90 which showed a reduction to 60% (Wiesler, Weinzierl 2011). The results further showed that recruitment of the RNA polymerase was reduced in case of I88A (35%), K89A (55%) and R90 (20%), whereas deletion of three amino acids K87 - K89 and I88 - R90 have no influence on transcription activity. However, the recruitment of the RNA polymerase seems not to be diminished in the results shown here, as preinitiation complexes were formed in the same amount as the wtTFB for all TFB variants used.

Nevertheless, amino acid E53 of *P. furiosus* TFB seems to be essential for transcription. Substitution of this amino acid with alanine leads to an altered DNA/TBP/TFB complex

pattern in EMSA experiments (Figure 17 A), and further showed almost no transcriptional activity at any assays. Only in experiments with pre-opened templates an activity of 49% (nt-strand mismatch template) and 68% (t-strand mismatch) was observed (Figure 21). The fact that the TBP/TFB/DNA complex runs lower on the gel in EMSAs suggests that the TFB conformation is disordered which results in a change of the electromobilic property. Less is known about the intramolecular interactions of TFB/TFIIB. Recent fluorescence studies of human TFIIB B-reader position W52 (*Pfu*TFB W44) suggested structural and dynamic changes of TFIIB after interaction with DNA, which further indicates that the conformation of TFIIB differs depending on the interactions to other components of the transcription machinery (Gorecki et al. 2015). It was also suggested from cryo-EM structures of the human transcription complex, that the B-reader domain is in a highly disordered conformation whereas tight interactions where formed when open complex formation takes place (Plaschka et al. 2016). Therefore E53 might be involved in maintaining a specific conformation of TFB if not part of the PIC, which would explain why TBP/TFB/DNA complexes showed altered electromobilic properties but are able to form initiation complexes. Interestingly, if the amino acid next to E53, R52, is also substituted with alanine, it stimulates formation of the first phosphodiester bond (Figure 18). The TFB variant R52E53A basically showed the best results in the screening in contrast to the single substitutions, R52A and E53A, suggesting that this charge distribution can stimulate transcription initiation and form run-off transcripts comparable to wtTFB. This finding leads to the conclusion that the acidic E53 and the basic R52 amino acid of the wild type can be replaced with two unpolar alanine amino acids at this site. Therefore an overall neutral charge distribution of these two positions is important for the function of the TFB B-reader loop. However, the LoopA mutation (which comprises four alanine from R52 to R55), and the double-substituted TFB-R54R55A showed complete loss of function in the transcription assays, even in presence of TFE. This finding indicates that the charge composition of the B-reader loop tip is of high importance for initiating transcription, especially for the conserved residue R55. Every alanine combination including this amino acid is not able for transcription at a sufficient level. This result leads to the suggestion that the two basic residues R54 and R55 are also important for the function of the B-reader loop. Taken together, this protein region tolerates the combinations AARR, AERR, and in little, RAAR to fulfill its role in transcription initiation. Interestingly, TFB variant R52Bpa of the crosslinking studies, which contains the negative charged unnatural phenylalanine derivate, showed 2-fold increase in abortive transcription assays (Figure 25 C). Therefore and the fact that corresponding eukaryotic TFIIB B-reader regions comprise an overall negative charge, it might be interesting if transcription can be enhanced by complete substitution of the archaeal TFB B-reader with acidic residues.

Taken together, the alanine screening of the TFB B-reader loop tip region showed that substitution of the wild type amino acids at distinct positions can result in a collapse of the transcription system, due to insufficient interactions of the selected protein region likely with the t-strand of the transcription bubble in proximity to the active site of the RNAP. It was shown that the acidic residue E53, together with the basic arginine R55 of the loop are necessary for correct transcription initiation, whereas the elimination of the overall charge of the loop tip is also insufficient for correct transcription initiation. Positive charged amino acids were stepwise eliminated, and the results suggest that important protein-DNA interactions strongly depend on the charge distribution of this region. Due to the close position of the TFB B-reader loop to the t-strand, it can be concluded that the overall positive charge of this region is important for correct DNA strand positioning and stabilization during the strand slips inside the cleft of the polymerase.

D.   RNA-strand separation does not depend on the charge of the B-reader loop

Despite the low ability of the TFB alanine substitutions to initiate transcription, RNA-strand separation was investigated. To overcome deficiencies in promoter opening and/or t-strand stabilization, a heteroduplex DNA template was used. Surprisingly, a three nucleotide mismatch (-1 to +2) was sufficient to restore almost wild type transcription levels for all TFB variants used (Figure 21). From this point of view it can be suggested that the TFB B-reader loop might be involved in promoter opening directly or indirectly. Due to its overall positive charge distribution, an interaction with the negative charged backbone might support strand separation or the slip of the single stranded DNA inside the cleft of the RNA polymerase, which can be an explanation for the fact that a mini bubble is sufficient for restoration of transcriptional activity. However, in experiments performed with the pre-opened template transcript formation took place at wtTFB level except E53A, which showed 50% activity at templates containing an nt-strand mismatch, and 68% activity at templates containing a t-strand mismatch. The difference observed at both templates further indicate a direct interaction of position E53 with the t-strand DNA, as transcription is increased to 140% in experiments performed with the template containing the mismatch at the t-strand. TFB-R54R55A and LoopA both showed a slight reduction in these experiments with values between 82 and 92% for both templates, respectively. Nevertheless, a defect in RNA strand separation can be excluded as run-off transcripts were formed at levels comparable to wtTFB. RNA was guided correctly towards the exit channel of the polymerase. If this event is not performed correctly, the number of formed transcripts shouldn't be at the wild type level. Therefore the charge of the B-reader loop domain does not influence the RNA-strand separation in *P. furiosus* like it was postulated for the related yeast transcription system.

E.   Topology of *Pfu*TFB is almost similar to TFIIB

As it was mentioned in the previous chapter, the B-reader domain is located in close proximity to the t-strand in an initially transcribing complex of a eukaryotic crystal structure (Sainsbury et al. 2013). A comparable structure of the related *P. furiosus* transcription system is missing, and structural information about the topology of archaeal TFB within the preinitiation complex was thought to be like in eukaryotic organisms, based on the relation of transcription components of the two domains. To reveal the position of the TFB domains within the complex crosslinking experiments were performed. In this approach site-specifically mutagenized TFB proteins, which contained the UV inducible crosslinking agent p-Benzoyl-L-phenylalanine (Bpa), were used together with site-specifically radiolabeled DNA templates. The TFB-Bpa variants of the B-reader domain (G41 - R57), the B-linker (E74 and M85) and the B-core (F192) were screened in different *in vitro* transcription assays to select applicable mutants based on the overall quality of the amino acids and their corresponding position within the initial transcribing complex structure. The results showed that 9 of 19 screened TFB variants were suitable for crosslinking, and six proteins, A46Bpa, R52Bpa, S56Bpa, E74Bpa, M85Bpa and F192Bpa were selected (Figure 28). In addition, six TFB-variants (G41Bpa, E43Bpa, R45Bpa, D48Bpa, Q51Bpa and R57Bpa) showed no transcriptional activity in the experiments, but formed preinitiation complexes. Substitutions of the corresponding amino acids in *M. jannaschii* revealed that the highly conserved residues *Mja*TFB-E78 (*Pfu*TFB E43), *Mja*TFB R80 (*Pfu*TFB R45) and *Mja*TFB R92 (*Pfu*TFB R57) with phenylalanine or a negative charge results in very low or loss of transcriptional activity, indicating that these residues are essential for the function of TFB (Wiesler, Weinzierl 2011). In case of *Mja*TFB R92 substitution with any amino acid results in loss of transcriptional

activity, indicating that the charge of this residue is essential. It was also reported for yeast TFIIB that *Sc*TFIIB R78 (*Pfu*TFB R57) is also essential for transcription (Bangur et al. 1997). In the yeast crystal structure of the initially transcribing complex, R78 was shown to interact with D70, F66 and G80 of TFIIB, which in turn interacts with the rudder element of the RNAP II, indicating that this residue is of high importance for the correct formation of Pol II-TFIIB contacts (Sainsbury et al. 2013). The results shown here together with results described in literature for TFB proteins of other organisms, it can be concluded that the conserved amino acids G41, E43, R45, and especially R57 are essential for the function of *Pfu*TFB during *P. furiosus* transcription.

The selected TFB variants were used in specific crosslinking experiments. The results showed that TFB mutations located within the B-reader domain (A46, R52 and S56) crosslinked to DNA which contained a radioactively labeled nucleotide at position -4t (Figure 32), whereas this contact was only observed in presence of RNAP, indicating an interaction to the transcribing strand specifically in the preinitiation complex. In contrast, control samples performed with the linker mutations E74Bpa and M85Bpa did not show a specific interaction. Renfrow et al. showed also an interaction between TFB and DNA at position -4t, which is almost absent in closed complexes, and strongly increased in reactions containing RNAP in the *Pfu* transcription system (Renfrow et al. 2004). Therefore the results presented here demonstrate that the B-reader domain specifically interact with the t-strand in open complexes only. The results further show that the B-reader domain is located next to the transcribing strand in proximity to the active site of the RNAP, which is in accordance to published structures of modelled open complex and initially transcribing complex (Kostrewa et al. 2009; Sainsbury et al. 2013) and cryo-EM structures of yeast and human preinitiation complexes (Plaschka et al. 2016). Therefore a similar topology of this TFB domain can be suggested for the archaeal *P. furiosus* transcription system.

In reactions using DNA templates labeled at -11t a crosslink reaction was observed for the linker position E74, but not for the B-reader positions, which emphasize the specificity of the used crosslinking method. E74 also crosslinked to the nt-strand labeled at position -8. Contacts between TFB and DNA at -10t but not at -12t (Renfrow et al. 2004), and at -9nt (Micorescu et al. 2008) were also reported in literature, and the results presented here demonstrated that E74 is likely the interacting residue and may explain the results observed in these studies. The interaction between E74 and the t-strand was suggested for the corresponding amino acid of *Sc*TFIIB in the open complex model (Kostrewa et al. 2009), whereas the corresponding *Pfu*TFB M85 amino acid contacts the nt-strand in the structure. However, PfuTFB M85Bpa does not contact one of both strands in the experiments, indicating that this residue is out of distance for a crosslink reaction. Therefore it can be suggested that the nt-strand is located closer to the B-linker strand than to the B-linker helix position M85, which differs from the open complex model, whereas the nt-strand is not clearly defined in this model, and is further not resolved in the initially transcribing complex. KMnO$_4$ footprints of the open region of *P. furiosus* transcription estimated the detectable size of the transcription bubble from -9 to +5 of the nt-strand (Spitalny, Thomm 2003). As E74Bpa crosslinks to both strands, it can be concluded that this amino acid position is in between of the two strands located at the upstream edge of the transcription bubble, possibly involved in stabilization of the separated strands. This finding is further in accordance with the published open complex model for eukaryotic TFIIB (Kostrewa et al. 2009) and with recent observations for the archaeal transcription system of *M. jannaschii* (Nagy et al. 2015).

In addition, the B-core position F192Bpa crosslinked specifically to DNA labeled at position -19t in the ternary DNA/TBP/TFB complex as well as in open complexes (Figure 33). The B-

core position F192 was selected for the experiments because a contact to DNA at -19t was proposed from a DNA/TBP/TFB-core crystal structure of the related *P. woesei* (Kosa et al. 1997). The results shown here demonstrated that this specific contact is also formed in the *P. furiosus* system, and might be the explanation why an interaction between TFB and DNA at -18t and -20t in open complexes and DNA/TBP/TFB complexes was also observed in the *Pfu* transcription system reported by Renfrow et al. 2004. However, the corresponding amino acid I209 of the yeast TFIIB showed also contacts to DNA in the open complex model (Kostrewa et al. 2009) but is located a few bases upstream. This finding can be explained with results of FRET measurements for the complete archaeal *M. jannaschii* preinitiation complex, where it was shown that archaeal TBP and TFB are located closer to the surface of the RNA polymerase (Nagy et al. 2015). This finding might explain the distance between I209 and -19t in the postulated eukaryotic structure, and would also suggest a closer location of TFB to RNAP in *Pyrococcus furiosus* transcription.

Taken together, the results of the crosslinking experiments in the preinitiation complex reveal an almost similar topology of the TFB B-reader and B-linker domain of the archaeal *P. furiosus* transcription in comparison to eukaryotic TFIIB in published structures, whereas the location of the nt-strand and the B-core position F192 seems to be slightly different.

F.   The TFB B-reader domain is displaced at register +10

The amino acid substitutions of the B-reader domain, TFB-A46Bpa, TFB-R52Bpa and TFB-S56Bpa were used in crosslinking experiments on stalled transcription complexes to reveal structural transitions of this domain during transcription initiation and transition to early elongation. The results showed that at register +6 the signal intensity of the obtained crosslink for position S56Bpa is decreased. Sainsbury et al. postulated a RNA-DNA separation model, which is based on the charge-specific interaction of the B-reader loop and the nascent RNA at this position (Sainsbury et al. 2013). In contrast, *P. furiosus* does not comprise a negative charged B-reader loop, but in accordance to the model, the homology between TFB and TFIIB, and the fact that the B-reader of *P. furiosus* is in a similar location than its eukaryotic counterpart, a direct interaction with RNA can be assumed. In the bacterial system, a similar contact between the σ3.2 region, which corresponds to the B-reader loop, and the nascent RNA was postulated using FRET technique, as the initial transcribing complex containing a 6mer RNA paused at this site (Duchi et al. 2016). Crosslinking experiments of human TFIIB in preinitiation complexes together with a three nucleotide long RNA further revealed a close location of TFIIB to RNA in proximity to the active site of the RNAP II (Bick et al. 2015), which also indicate a TFB-RNA interaction observed in the experiments. A further interaction between the B-reader domain and the RNA was detected in experiments at register +8 for the helix position A46 (Figure 35). The signal decreased to approximately 70%, indicating a larger distance between the DNA and the incorporated Bpa. RNA was predicted to clash with the helix of TFIIB at positon +8 in the open complex model (Kostrewa et al. 2009). The results presented here confirmed the postulated interaction between RNA and the B-reader helix. The biggest changes in the crosslink pattern were observed in register +10 for all TFB B-reader variants. Here signal intensities dropped down to approximately 30% in comparison to the signals of the preinitiation complex and remained unchanged low at register +15, indicating a structural transition of the complete B-reader domain. The use of the B-core position F192 in crosslinking experiments on stalled complexes at this register demonstrated that TFB is fully present, indicating that the observed changes for the B-reader domain are specific for this protein region. Pal et al. postulated a transition of the corresponding TFIIB domain due to the

collapse of the transcription bubble at registers +10/+11 (Pal et al. 2005). In the bacterial system a transition of the corresponding region σ3.2 was observed, as this domain is in path of the advancing 5′end of the RNA, but takes place at register +6/+7 (Basu et al. 2014). In addition, $KMnO_4$ footprint experiments also reported a reduction in the size of the transcription bubble at registers +10/+11 for the *P. furiosus* transcription system (Spitalny, Thomm 2003). In accordance with the herein obtained results, the TFB B-reader domain is displaced at register +10, possibly due to interactions with RNA and the B-reader helix. The B-reader displacement further results in the collapse of the transcription bubble, as the stabilizing interaction between DNA and B-reader is attenuated.

G. TFB tends to be released from register +15 onwards

Using the TFB-F192Bpa variant the point of TFB release was determined. The results show that the signal intensities of TFB-DNA interactions in stalled complexes at registers +6 to +14 do not change in comparison to the signals derived from the respective preinitiation complexes, indicating that TFB is not released at these positions. At register +15 the signal is slightly decreased, and further at +20, whereas crosslinking experiments under run-off conditions showed just background signals. The results indicate that TFB release starts from +14 onwards, but is not an instantaneous event as it was expected. For the human transcription system a TFIIB release was described to take place at an RNA length of 12/13nt *in vitro* (Cabart et al. 2011). It was also postulated that the matured RNA with the same length clashes with the Zn-ribbon domain of TFIIB in the crystal structure of an initially transcribing complex of yeast, as this domain blocks the exit pore of the RNA polymerase II (Sainsbury et al. 2013). It was also shown by Xie et al that TFB is released in *in vitro* experiments lacking TFE in the archaeal M. *thermoautotrophicus* system when complexes were chased to position +24 (Xie, Reeve 2004b), as well as for the human *in vitro* transcription system, for which a TFIIB release was shown to take place between +6 and +16 (Tran, Gralla 2008). However, a release of TFB in the *P. furiosus* transcription was expected at registers +12/+13, but this event could not be pinpointed to these distinct nucleotide positions. However, due to the fact that the signal started to decrease at register +15 onwards, it can be suggested that TFB is destabilized at registers +13/+14 and starts to be released at register +15, but it remains speculative why the transcription factor is not released completely at one distinct register. It is possible that the release process happens slowly due to persisting interaction between TFB and DNA, TBP and RNAP.

H. Concluding aspects

Archaeal TFB and eukaryotic TFIIB show a high degree of structural and functional conservation. Both proteins are very important for transcription, as both factors fulfill crucial steps during initiation. In *P. furiosus*, TFB is sufficient to form a stable DNA/TBP/TFB complex with DNA in a bent conformation. The B-reader loop of *Pfu*TFB comprises a positive charge, but the corresponding region of the eukaryotic *Sc*TFIIB is negative. The charge distribution of this domain is essential for the function of *Pfu*TFB during transcription likely to stabilize the transcribed strand of the transcription bubble. In contrast, this domain is not involved in RNA-strand separation, like it was proposed for eukaryotic TFIIB. In addition, archaeal preinitiation complexes consist of a reduced number of transcription factors than the eukaryotic complex, but it has to undergo nearly the same transitions to initiate RNA synthesis. For example, the TFIIB/TFIID stabilizing factor TFIIA, as well as TFIIH, which supports DNA melting and translocation of the RNAP are missing in archaeal transcription. Therefore, archaeal transcription is a simplified version of the eukaryotic transcription, and is more deeply rooted in the tree of life. Despite similar roles in both transcription systems, the
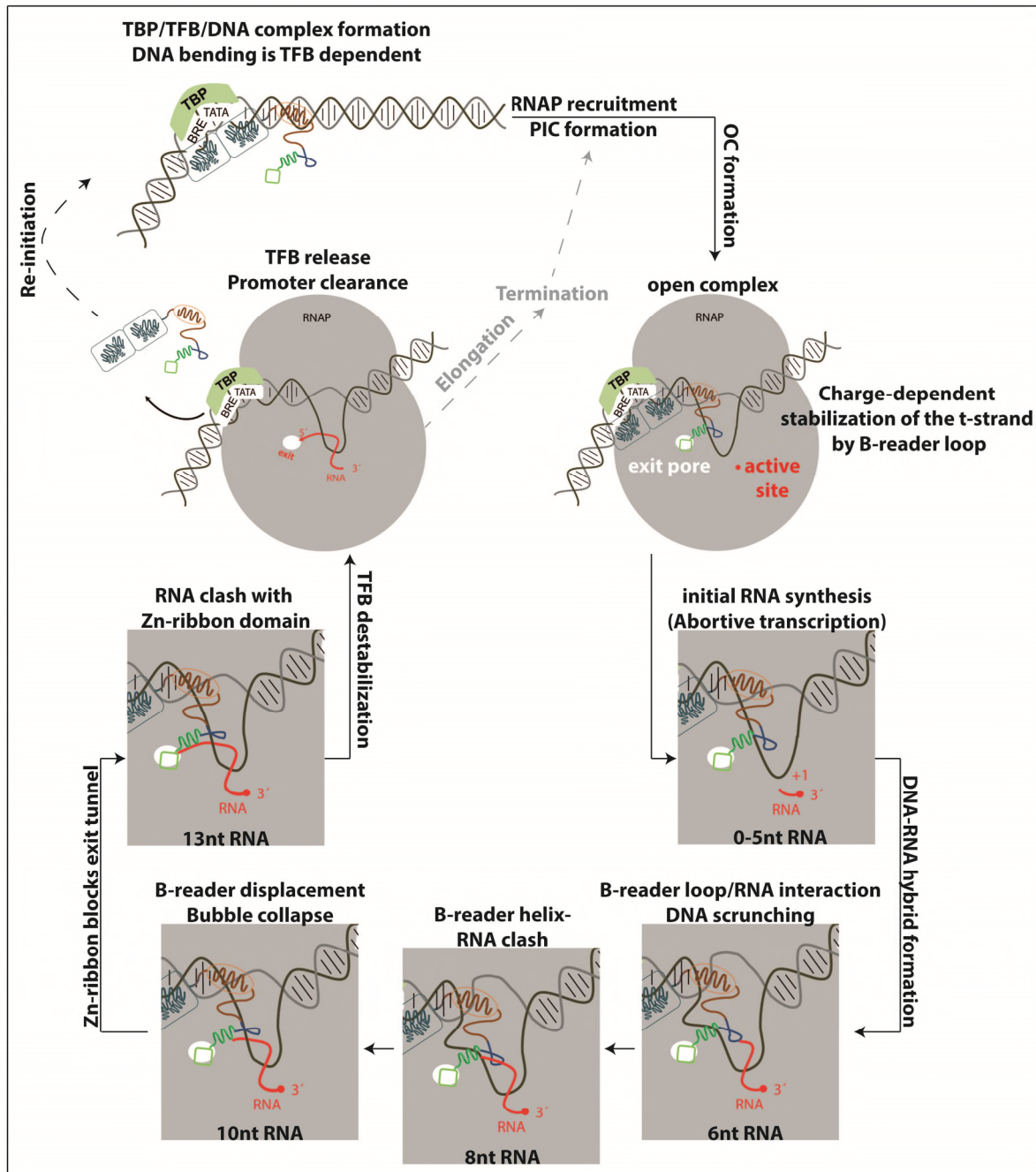
Figure 38: Complete initiation of transcription of *Pyrococcus furiosus*. Transcription starts with binding of TBP and TFB to the promoter site, whereas DNA bending relies on the presence of TFB. The polymerase is recruited to the promoter to form a preinitiation complex. DNA is melted around the start site whereas the charge of the TFB B-reader loop region is important for t-strand loading and stabilization of the single-stranded area. The RNA polymerase starts to synthesize RNA, and at a length of 6nt it interacts with the TFB B-reader loop region, whereas the charge of this loop is not required for RNA-strand separation. RNA then clashes with TFB B-reader helix at a length of 8nt, and at a length of 10nt it displaces the TFB B-reader domain. This translocation results in collapse of the transcription bubble, and in a further destabilization of TFB. The transcription factor is completely destabilized at register +13/+14 possibly due to a clash of RNA with the Zn-ribbon, which induces the release of TFB from register +15 onwards. The RNA polymerase can reinitiate to the promoter to start the next transcription cycle.

archaeal TFB might have more fundamental roles during transcription, as the eukaryotic system gained several additional factors which are involved in initiating processes.

The role of the transcription factor B in transcription initiation of *P. furiosus* with respect to the results obtained in this thesis is summarized in figure 38.

Despite the reduced degree of complexity of the archaeal transcription system, several postulations derived from eukaryotic cryo-EM and crystal structures were addressed in this thesis using crosslinking experiments in stalled transcription complexes. The crosslinking method used in this study was highly specific and enabled detection of TFB-DNA contacts during transition from initiation to elongation and monitoring dynamic transitions of TFB. In addition to other methods like single molecule FRET analysis, crosslinking is also a powerful tool to elucidate functional interactions and structural rearrangements of proteins and their targets. Crosslinking might be applicable to investigate DNA scrunching during transcription initiation, as well as interactions of other DNA binding proteins, which are involved in regulation of transcription.

The results of the crosslinking experiments presented here show the first dynamic transitions of the archaeal transcription factor B, and provide evidence for structural rearrangements within the complex during transition from initiation to elongation. The data further confirm several postulated events derived from eukaryotic complexes and therefore complement the structural information on a biochemical level. The results further give a better understanding of the archaeal transcription initiation mechanism and show similarities as well as differences between the archaeal and the eukaryotic transcription machineries.

VI.   Abstract

The preinitiation complex of the transcription machinery in archaeal organisms resembles a simplified version of the eukaryotic RNA polymerase II transcription system. Both systems share homologous general transcription factors to recruit RNA polymerase to the promoter to initiate RNA synthesis. The transcription factor (II)B plays an important role during transcription initiation. Based on eukaryotic cryo-EM and crystal structures several functional interactions and structural transitions of TF(II)B were proposed. To detect specific interactions of the archaeal *P. furiosus* TFB during transcription initiation different *in vitro* transcription assays were performed. In addition, the replication protein A of *P. furiosus* was also investigated using various *in vitro* experiments.

Crosslinking experiments using TFB, which contained a UV inducible photo crosslinker, and site-specific radioactively labeled DNA templates revealed an almost similar topology of the archaeal TFB B-reader and B-linker domains in the preinitiation complex in comparison to corresponding regions predicted in eukaryotic structures. Unlike it was postulated in open complex models, the non-transcribed strand is located closer to the B-linker strand than the B-linker helix. The B-core amino acid F192 contact DNA 19 nucleotides upstream the transcribed strand, in accordance to a published crystal of *P. woesei* TATA/TBP/TFB-core structure, but is different to predicted eukaryotic closed and open complex models. Crosslinking experiments in stalled complexes showed that RNA interacts with the B-reader loop at a length of 6nt, and further clashes with the B-reader helix domain with a length of 8nt. At register +10 the TFB B-reader is displaced, which causes collapse of the transcription bubble. It was also demonstrated that TFB is present at register +6 to +14 in the complex, and tended to be released from register +15 onwards, indicating a destabilization of TFB at register +13/+14.

Alanine substitutions of amino acids of the TFB B-reader loop revealed that this region mainly stabilizes the transcription bubble due to charge-dependent interactions with the transcribing strand. In contrast to the predicted RNA-DNA separation model derived from a eukaryotic initially transcribing complex, RNA-strand separation does not depend on the charge of the *Pfu*TFB B-reader loop.

Single molecule FRET experiments revealed that DNA bending depends on the presence of TFB in *P. furiosus*.

In vitro transcription assays with RPA showed that this protein has binding preference to single stranded DNA. Experiments further showed that RPA is not involved in transcription initiation, but it stimulates transcription. Therefore RPA functions during elongation of transcription, possibly due to a stabilization of the RNA polymerase and increase of the processivity.

The results presented here give a more detailed insight into molecular interactions of TFB and are the first biochemical data on dynamic rearrangements of TFB during transcription initiation and transition to early elongation. It further deepens the understanding of archaeal transcription processes and complements structural information derived from related eukaryotic organisms.

VII. Zusammenfassung

Der Präinitiationskomplex der Transkriptionsmaschinerie archaeeller Organismen gleicht einer vereinfachten Version des eukaryotischen RNS Polymerase II Komplexes. Beide Systeme verwenden zum Teil homologe Transkriptionsfaktoren, um die RNS Polymerase zur Initiierung der RNS-Synthese an den Promoter zu rekrutieren. Der Transkriptionsfaktor (II)B hat dabei mehrere wichtige Funktionen in diesem Komplex. Basierend auf eukaryotischen Kristall- und cryo-EM Strukturen wurden funktionelle Interaktionen und strukturelle Veränderungen vorhergesagt. Um spezifische Wechselwirkungen des archaeellen TFB aus *P. furiosus* während der Transkriptionsinitiation zu detektieren, wurden verschiedene *in vitro* Transkriptionsexperimente verwendet. Zudem wurde ein weiteres Protein, das Replikationsprotein A aus *P. furiosus*, in verschiedenen Experimenten hinsichtlich dessen Funktion untersucht.

Crosslink-Experimente, in denen TFB mit einem UV-induzierbaren Crosslinker ausgestattet und zusammen mit spezifisch radioaktiv markierten DNS Matrizen verwendet wurde, zeigten, dass die TFB B-reader und B-linker Domänen eine nahezu ähnliche Lage im Präinitiationskomplex aufweisen, wie die entsprechenden Domänen in eukaryotischen Strukturen. Anders als im Model eines offenen Transkriptionskomplexes liegt der nicht-transkribierte Strang näher am B-linker Strang als an der B-linker Helix. In Übereinstimmung mit einer publizierten *P. woesei* TATA/TBP/TFB-Kern Struktur zeigte die Aminosäure F192 aus der TFB Kerndomäne einen Kontakt zur DNS 19 Nukleotide stromaufwärts am transkribierten Strang, und weist damit Unterschiede zu modellierten Strukturen von geschlossenen und offenen eukaryotischen Transkriptionskomplexen auf. Crosslink-Experimente in gestellten Komplexen zeigten, das RNA mit einer Länge von 6nt mit der B-reader Schleife interagiert, und anschließend mit einer Länge von 8nt mit der B-reader Helix zusammenstößt. An Position +10 ist die TFB B-reader Domäne verschoben, was zu einem Zusammenbruch der Transkriptionsblase führt. Es konnte auch gezeigt werden, dass TFB in den Registern +6 bis +14 im Komplex vorhanden ist, und dazu tendiert, ab Position +15 freigesetzt zu werden, was auf eine Destabilisierung des TFB an Position +13/+14 hindeutet.

Alanin-Substitutionen von Aminosäuren der TFB B-reader Schleife zeigten, dass diese Region hauptsächlich die Transkriptionsblase stabilisiert aufgrund ladungsabhängiger Wechselwirkungen mit dem transkribierten Strang. Im Gegensatz zum postulierten RNS-DNS Separationsmodel basierend auf einer eukaryotischen Struktur eines initial transkribierenden Komplexes ist die Trennung der Stränge nicht von der Ladung der TFB B-reader Schleife abhängig.

In Einzelmolekül-FRET Studien konnte gezeigt werden, dass die DNS-Biegung in *P. furiosus* von TFB abhängig ist.

In vitro Studien mit RPA zeigten, dass dieses Protein eine Einzelstrang-Präferenz besitzt. Die Experimente deuten darauf hin, dass RPA nicht an der Initiation beteiligt ist, aber Transkription konnte stimuliert werden. Es wurde gezeigt, dass RPA in die Elongation der Transkription eingreift, da es möglicherweise die RNS Polymerase stabilisiert und dessen Prozessivität erhöht.

Die hier gezeigten Daten geben einen detaillierteren Einblick in molekulare Interaktionen von TFB, und sind die ersten biochemischen Daten über dynamische Veränderungen von TFB während der Transkriptionsinitiation und dem Übergang in die frühe Elongation. Das Verständnis der archaeellen Transkriptionsprozesse soll damit vertieft werden, und die strukturellen Informationen aus den verwandten eukaryotischen Organismen komplettieren.

VIII. Appendix

A. Abbreviation list

α	Alpha
A	Ampere
A [%]	Average in percentage
Abs	Absorption
AGE	Agarose gel electrophoresis
Å	Angstrom
ATP	Adenosine triphosphate
β	Beta
b	Base
bp	Base pair
Bq	Becquerel
Bpa	p-Benzoyl-*L*-phenylalanine
BSA	Bovine serum albumin
c	Centi
°C	Degree Celsius
CPM	Counts per minute
C-term.	Carboxy terminal
CTP	Cytosine triphosphate
Da	Dalton
Δ	Delta
dIC	Poly 2´deoxyinosinic-2´-deoxycytidylic acid
DNA	Deoxyribonucleic acid
dNTP	deoxy nucleoside triphosphate
ds	Double stranded
DTT	1,4 Dithiothreitol
E.	Escherichia
EDTA	Ethylenediaminetetraacetic acid
e.g.	exempli gratia
EM	Electron microscopy
Em	Emission
EMSA	Electro mobility shift assay
et al.	et alii
f	Femto
F	Forward
FAM	6-Carboxyfluorescein
FRET	Förster resonance energy transfer
γ	Gamma
g	Gram
g (=RCF)	Relative centrifugal force
gdh	Glutamate dehydrogenase
GpU	Guanylyl-5´-phosphatidyl-Uracil
GTF	General transcription factor
GTP	Guanine triphosphate
IPTG	Isopropyl β-D-1-thiogalactopyranoside
k	Kilo
λ	Lambda
l	Liter
LB	Lysogeny broth
μ	Micro
m	Milli
M	Mega
M (chem. unit)	Molar
Mja	*Methanocaldococcus jannaschii*
n	Nano
NaAc	Sodium acetate
nt	Nucleotide

N-term.	Amino terminal
NTP	Nucleoside triphosphate
nt-strand	Non-transcribing strand
OD	Optical density
p	Pico
p.a.	pro analysi
Pab	*Pyrococcus abyssi*
PAGE	Poly acrylamide gel electrophoresis
pH	Pondus hydrogenii
PCI	Phenol/Chloroform Isoamyl alcohol
PCR	Polymerase chain reaction
Pfu	*Pyrococcus furiosus*
PIC	Preinitiation complex
PMSF	Phenylmethylsulfonyl fluoride
Pol	Polymerase
PVDF	Polyvinylidene fluoride
R	Reverse
RNA	Ribonucleic acid
RNAP	RNA polymerase
RPA	Replication protein A
s (sec)	Second
σ	Sigma
Sc	*Saccharomyces cerevisiae*
SD [%]	Standard deviation in percentage
SDS	Sodium dodecyl sulfate
ss	Single stranded
SSB	Single stranded binding
T	Tera
TBP	TATA binding protein
TEC	Ternary elongation complex
TEMED	Tetramethylethylenediamine
TF	Transcription factor
TFB	Transcription factor B
TFE	Transcription factor E
TIRF	Total internal reflection fluorescence
t-strand	Transcribing strand
TSS	Transcription start site
U	Unit
UTP	Uracil triphosphate
UV	Ultra violet
V	Volt
v/v	Volume per volume
W	Watt
wt	Wild type
w/v	Weight per volume

B.   Figure list

IX. Publication bibliography

Abramoff, M.D., Magalhaes, P.J., Ram, S.J: Image Processing with ImageJ. In *Biophotonics International* 2004 (Volume 11, Issue 7), pp. 36–42.

Adachi, Naruhiko; Senda, Miki; Natsume, Ryo; Senda, Toshiya; Horikoshi, Masami (2008): Crystal structure of Methanococcus jannaschii TATA box-binding protein. In *Genes to cells : devoted to molecular & cellular mechanisms* 13 (11), pp. 1127–1140. DOI: 10.1111/j.1365-2443.2008.01233.x.

Akhtar, Waseem; Veenstra, Gert Jan C. (2011): TBP-related factors: a paradigm of diversity in transcription initiation. In *Cell & bioscience* 1 (1), p. 23. DOI: 10.1186/2045-3701-1-23.

Arimbasseri, Aneeshkumar G.; Maraia, Richard J. (2016): RNA Polymerase III Advances: Structural and tRNA Functional Views. In *Trends in biochemical sciences* 41 (6), pp. 546–559. DOI: 10.1016/j.tibs.2016.03.003.

Avery, O. T.; Macleod, C. M.; McCarty, M. (1944): STUDIES ON THE CHEMICAL NATURE OF THE SUBSTANCE INDUCING TRANSFORMATION OF PNEUMOCOCCAL TYPES. INDUCTION OF TRANSFORMATION BY A DESOXYRIBONUCLEIC ACID FRACTION ISOLATED FROM PNEUMOCOCCUS TYPE III. In *The Journal of experimental medicine* 79 (2), pp. 137–158.

Bae, Brian; Chen, James; Davis, Elizabeth; Leon, Katherine; Darst, Seth A.; Campbell, Elizabeth A. (2015): CarD uses a minor groove wedge mechanism to stabilize the RNA polymerase open promoter complex. In *eLife* 4. DOI: 10.7554/eLife.08505.

Bagby, S.; Kim, S.; Maldonado, E.; Tong, K. I.; Reinberg, D.; Ikura, M. (1995): Solution structure of the C-terminal core domain of human TFIIB: similarity to cyclin A and interaction with TATA-binding protein. In *Cell* 82 (5), pp. 857–867.

Bajic, Vladimir B.; Tan, Sin Lam; Christoffels, Alan; Schonbach, Christian; Lipovich, Leonard; Yang, Liang et al. (2006): Mice and men: their promoter properties. In *PLoS genetics* 2 (4), e54. DOI: 10.1371/journal.pgen.0020054.

Bangur, C. S.; Pardee, T. S.; Ponticelli, A. S. (1997): Mutational analysis of the D1/E1 core helices and the conserved N-terminal region of yeast transcription factor IIB (TFIIB): identification of an N-terminal mutant that stabilizes TATA-binding protein-TFIIB-DNA complexes. In *Molecular and cellular biology* 17 (12), pp. 6784–6793.

Basu, Ritwika S.; Warner, Brittany A.; Molodtsov, Vadim; Pupov, Danil; Esyunina, Daria; Fernandez-Tornero, Carlos et al. (2014): Structural basis of transcription initiation by bacterial RNA polymerase holoenzyme. In *The Journal of biological chemistry* 289 (35), pp. 24549–24559. DOI: 10.1074/jbc.M114.584037.

Bell, S. D.; Brinkman, A. B.; van der Oost, J.; Jackson, S. P. (2001): The archaeal TFIIEalpha homologue facilitates transcription initiation by enhancing TATA-box recognition. In *EMBO reports* 2 (2), pp. 133–138. DOI: 10.1093/embo-reports/kve021.

Bell, S. D.; Jackson, S. P. (1998): Transcription in Archaea. In *Cold Spring Harbor symposia on quantitative biology* 63, pp. 41–51.

Bell, S. D.; Jackson, S. P. (2001): Mechanism and regulation of transcription in archaea. In *Current opinion in microbiology* 4 (2), pp. 208–213.

Bell, S. D.; Jaxel, C.; Nadal, M.; Kosa, P. F.; Jackson, S. P. (1998): Temperature, template topology, and factor requirements of archaeal transcription. In *Proceedings of the National Academy of Sciences of the United States of America* 95 (26), pp. 15218–15222.

Bell, S. D.; Kosa, P. L.; Sigler, P. B.; Jackson, S. P. (1999): Orientation of the transcription preinitiation complex in archaea. In *Proceedings of the National Academy of Sciences of the United States of America* 96 (24), pp. 13662–13667.

Bhartiya, Deeksha; Scaria, Vinod (2016): Genomic variations in non-coding RNAs: Structure, function and regulation. In *Genomics* 107 (2-3), pp. 59–68. DOI: 10.1016/j.ygeno.2016.01.005.

Bick, Matthew J.; Malik, Sohail; Mustaev, Arkady; Darst, Seth A. (2015): TFIIB is only approximately 9 A away from the 5'-end of a trimeric RNA primer in a functional RNA polymerase II preinitiation complex. In *PloS one* 10 (3), e0119007. DOI: 10.1371/journal.pone.0119007.

Bleiholder, Anne; Frommherz, Regina; Teufel, Katharina; Pfeifer, Felicitas (2012): Expression of multiple tfb genes in different Halobacterium salinarum strains and interaction of TFB with transcriptional activator GvpE. In *Archives of microbiology* 194 (4), pp. 269–279. DOI: 10.1007/s00203-011-0756-z.

Blombach, Fabian; Salvadori, Enrico; Fouqueau, Thomas; Yan, Jun; Reimann, Julia; Sheppard, Carol et al. (2015): Archaeal TFEalpha/beta is a hybrid of TFIIE and the RNA polymerase III subcomplex hRPC62/39. In *eLife* 4, e08378. DOI: 10.7554/eLife.08378.

Blombach, Fabian; Smollett, Katherine L.; Grohmann, Dina; Werner, Finn (2016): Molecular Mechanisms of Transcription Initiation-Structure, Function, and Evolution of TFE/TFIIE-Like Factors and Open Complex Formation. In *Journal of molecular biology* 428 (12), pp. 2592–2606. DOI: 10.1016/j.jmb.2016.04.016.

Bochkareva, E.; Korolev, S.; Bochkarev, A. (2000): The role for zinc in replication protein A. In *The Journal of biological chemistry* 275 (35), pp. 27332–27338. DOI: 10.1074/jbc.M000620200.

Bradford, M. M. (1976): A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. In *Analytical biochemistry* 72, pp. 248–254.

Bratkovic, Tomaz; Rogelj, Boris (2014): The many faces of small nucleolar RNAs. In *Biochimica et biophysica acta* 1839 (6), pp. 438–443. DOI: 10.1016/j.bbagrm.2014.04.009.

Brindefalk, Bjorn; Dessailly, Benoit H.; Yeats, Corin; Orengo, Christine; Werner, Finn; Poole, Anthony M. (2013): Evolutionary history of the TBP-domain superfamily. In *Nucleic acids research* 41 (5), pp. 2832–2845. DOI: 10.1093/nar/gkt045.

Brochier-Armanet, Celine; Forterre, Patrick (2007): Widespread distribution of archaeal reverse gyrase in thermophilic bacteria suggests a complex history of vertical inheritance and lateral gene transfers. In *Archaea (Vancouver, B.C.)* 2 (2), pp. 83–93.

Brueckner, Florian; Ortiz, Julio; Cramer, Patrick (2009): A movie of the RNA polymerase nucleotide addition cycle. In *Current opinion in structural biology* 19 (3), pp. 294–299. DOI: 10.1016/j.sbi.2009.04.005.

Brun, I.; Sentenac, A.; Werner, M. (1997): Dual role of the C34 subunit of RNA polymerase III in transcription initiation. In *The EMBO journal* 16 (18), pp. 5730–5741. DOI: 10.1093/emboj/16.18.5730.

Buratowski, S.; Hahn, S.; Guarente, L.; Sharp, P. A. (1989): Five intermediate complexes in transcription initiation by RNA polymerase II. In *Cell* 56 (4), pp. 549–561.

Buratowski, S.; Zhou, H. (1993): Functional domains of transcription factor TFIIB. In *Proceedings of the National Academy of Sciences of the United States of America* 90 (12), pp. 5633–5637.

Burke, T. W.; Kadonaga, J. T. (1996): Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. In *Genes & development* 10 (6), pp. 711–724.

Burley, S. K. (1996): The TATA box binding protein. In *Current opinion in structural biology* 6 (1), pp. 69–75.

Burley, S. K.; Roeder, R. G. (1996): Biochemistry and structural biology of transcription factor IID (TFIID). In *Annual review of biochemistry* 65, pp. 769–799. DOI: 10.1146/annurev.bi.65.070196.004005.

Burton, Samuel P.; Burton, Zachary F. (2014): The sigma enigma: bacterial sigma factors, archaeal TFB and eukaryotic TFIIB are homologs. In *Transcription* 5 (4), e967599. DOI: 10.4161/21541264.2014.967599.

Burton, Z. F.; Ortolan, L. G.; Greenblatt, J. (1986): Proteins that bind to RNA polymerase II are required for accurate initiation of transcription at the adenovirus 2 major late promoter. In *The EMBO journal* 5 (11), pp. 2923–2930.

Bushnell, D. A.; Bamdad, C.; Kornberg, R. D. (1996): A minimal set of RNA polymerase II transcription protein interactions. In *The Journal of biological chemistry* 271 (33), pp. 20170–20174.

Bushnell, David A.; Westover, Kenneth D.; Davis, Ralph E.; Kornberg, Roger D. (2004): Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms. In *Science (New York, N.Y.)* 303 (5660), pp. 983–988. DOI: 10.1126/science.1090838.

Butler, Jennifer E. F.; Kadonaga, James T. (2002): The RNA polymerase II core promoter: a key component in the regulation of gene expression. In *Genes & development* 16 (20), pp. 2583–2592. DOI: 10.1101/gad.1026202.

Cabart, Pavel; Ujvari, Andrea; Pal, Mahadeb; Luse, Donal S. (2011): Transcription factor TFIIF is not required for initiation by RNA polymerase II, but it is essential to stabilize transcription factor TFIIB in early elongation complexes. In *Proceedings of the National Academy of Sciences of the United States of America* 108 (38), pp. 15786–15791. DOI: 10.1073/pnas.1104591108.

Callaci, S.; Heyduk, E.; Heyduk, T. (1998): Conformational changes of Escherichia coli RNA polymerase sigma70 factor induced by binding to the core enzyme. In *The Journal of biological chemistry* 273 (49), pp. 32995–33001.

Carlo, Sacha de; Lin, Shih-Chieh; Taatjes, Dylan J.; Hoenger, Andreas (2010): Molecular basis of transcription initiation in Archaea. In *Transcription* 1 (2), pp. 103–111. DOI: 10.4161/trns.1.2.13189.

Carninci, Piero; Sandelin, Albin; Lenhard, Boris; Katayama, Shintaro; Shimokawa, Kazuro; Ponjavic, Jasmina et al. (2006): Genome-wide analysis of mammalian promoter architecture and evolution. In *Nature genetics* 38 (6), pp. 626–635. DOI: 10.1038/ng1789.

Carpousis, A. J.; Gralla, J. D. (1980): Cycling of ribonucleic acid polymerase to produce oligonucleotides during initiation in vitro at the lac UV5 promoter. In *Biochemistry* 19 (14), pp. 3245–3253.

Carter, Robert; Drouin, Guy (2010): The increase in the number of subunits in eukaryotic RNA polymerase III relative to RNA polymerase II is due to the permanent recruitment of general transcription factors. In *Molecular biology and evolution* 27 (5), pp. 1035–1043. DOI: 10.1093/molbev/msp316.

Chafin, D. R.; Claussen, T. J.; Price, D. H. (1991): Identification and purification of a yeast protein that affects elongation by RNA polymerase II. In *The Journal of biological chemistry* 266 (14), pp. 9256–9262.

Chakraborty, Anirban; Wang, Dongye; Ebright, Yon W.; Korlann, You; Kortkhonjia, Ekaterine; Kim, Taiho et al. (2012): Opening and closing of the bacterial RNA polymerase clamp. In *Science (New York, N.Y.)* 337 (6094), pp. 591–595. DOI: 10.1126/science.1218716.

Chasman, D. I.; Flaherty, K. M.; Sharp, P. A.; Kornberg, R. D. (1993): Crystal structure of yeast TATA-binding protein and model for interaction with DNA. In *Proceedings of the National Academy of Sciences of the United States of America* 90 (17), pp. 8174–8178.

Chedin, S.; Riva, M.; Schultz, P.; Sentenac, A.; Carles, C. (1998): The RNA cleavage activity of RNA polymerase III is mediated by an essential TFIIS-like subunit and is important for transcription termination. In *Genes & development* 12 (24), pp. 3857–3871.

Chen, Hung-Ta; Hahn, Steven (2003): Binding of TFIIB to RNA polymerase II: Mapping the binding site for the TFIIB zinc ribbon domain within the preinitiation complex. In *Molecular cell* 12 (2), pp. 437–447.

Cheung, P.; Allis, C. D.; Sassone-Corsi, P. (2000): Signaling to chromatin through histone modifications. In *Cell* 103 (2), pp. 263–271.

Cho, Byung-Kwan; Barrett, Christian L.; Knight, Eric M.; Park, Young Seoub; Palsson, Bernhard O. (2008): Genome-scale reconstruction of the Lrp regulatory network in Escherichia coli. In *Proceedings of the National Academy of Sciences of the United States of America* 105 (49), pp. 19462–19467. DOI: 10.1073/pnas.0807227105.

Cho, E. J.; Buratowski, S. (1999): Evidence that transcription factor IIB is required for a post-assembly step in transcription initiation. In *The Journal of biological chemistry* 274 (36), pp. 25807–25813.

Coker, James A.; DasSarma, Shiladitya (2007): Genetic and transcriptomic analysis of transcription factor genes in the model halophilic Archaeon: coordinate action of TbpD and TfbA. In *BMC genetics* 8, p. 61. DOI: 10.1186/ 1471-2156-8-61.

Colgan, J.; Manley, J. L. (1995): Cooperation between core promoter elements influences transcriptional activity in vivo. In *Proceedings of the National Academy of Sciences of the United States of America* 92 (6), pp. 1955–1959.

Cramer, P.; Bushnell, D. A.; Fu, J.; Gnatt, A. L.; Maier-Davis, B.; Thompson, N. E. et al. (2000): Architecture of RNA polymerase II and implications for the transcription mechanism. In *Science (New York, N.Y.)* 288 (5466), pp. 640–649.

Crick, F. (1958): On protein synthesis. In *Symposia of the Society for Experimental Biology* 12, pp. 138–163.

Crick, F. (1970): Central dogma of molecular biology. In *Nature* 227 (5258), pp. 561–563.

Cubonova, L'ubomira; Sandman, Kathleen; Hallam, Steven J.; Delong, Edward F.; Reeve, John N. (2005): Histones in crenarchaea. In *Journal of bacteriology* 187 (15), pp. 5482–5485. DOI: 10.1128/JB.187.15.5482-5485.2005.

Dahlke, Isabell; Thomm, Michael (2002): A Pyrococcus homolog of the leucine-responsive regulatory protein, LrpA, inhibits transcription by abrogating RNA polymerase recruitment. In *Nucleic acids research* 30 (3), pp. 701–710.

Davis, Elizabeth; Chen, James; Leon, Katherine; Darst, Seth A.; Campbell, Elizabeth A. (2015): Mycobacterial RNA polymerase forms unstable open promoter complexes that are stabilized by CarD. In *Nucleic acids research* 43 (1), pp. 433–445. DOI: 10.1093/nar/gku1231.

Decker, Kimberly B.; Hinton, Deborah M. (2013): Transcription regulation at the core: similarities among bacterial, archaeal, and eukaryotic RNA polymerases. In *Annual review of microbiology* 67, pp. 113–139. DOI: 10.1146/annurev-micro-092412-155756.

Demeny, Mate A.; Soutoglou, Evi; Nagy, Zita; Scheer, Elisabeth; Janoshazi, Agnes; Richardot, Magalie et al. (2007): Identification of a small TAF complex and its role in the assembly of TAF-containing complexes. In *PloS one* 2 (3), e316. DOI: 10.1371/journal.pone.0000316.

Deng, Wensheng; Roberts, Stefan G. E. (2005): A core promoter element downstream of the TATA box that is recognized by TFIIB. In *Genes & development* 19 (20), pp. 2418–2423. DOI: 10.1101/gad.342405.

Dieci, G.; Percudani, R.; Giuliodori, S.; Bottarelli, L.; Ottonello, S. (2000): TFIIIC-independent in vitro transcription of yeast tRNA genes. In *Journal of molecular biology* 299 (3), pp. 601–613. DOI: 10.1006/jmbi.2000.3783.

Dieci, Giorgio; Bosio, Maria Cristina; Fermi, Beatrice; Ferrari, Roberto (2013): Transcription reinitiation by RNA polymerase III. In *Biochimica et biophysica acta* 1829 (3-4), pp. 331–341. DOI: 10.1016/j.bbagrm.2012.10.009.

Dorman, Charles J. (2009): Nucleoid-associated proteins and bacterial physiology. In *Advances in applied microbiology* 67, pp. 47–64. DOI: 10.1016/S0065-2164(08)01002-2.

Dorman, Charles J. (2014): Function of nucleoid-associated proteins in chromosome structuring and transcriptional regulation. In *Journal of molecular microbiology and biotechnology* 24 (5-6), pp. 316–331. DOI: 10.1159/000368850.

Dorman, G.; Prestwich, G. D. (1994): Benzophenone photophores in biochemistry. In *Biochemistry* 33 (19), pp. 5661–5673.

Driessen, Rosalie P. C.; Dame, Remus Th (2011): Nucleoid-associated proteins in Crenarchaea. In *Biochemical Society transactions* 39 (1), pp. 116–121. DOI: 10.1042/BST0390116.

Drygin, Denis; Rice, William G.; Grummt, Ingrid (2010): The RNA polymerase I transcription machinery: an emerging target for the treatment of cancer. In *Annual review of pharmacology and toxicology* 50, pp. 131–156. DOI: 10.1146/annurev.pharmtox.010909.105844.

Duchi, Diego; Bauer, David L. V.; Fernandez, Laurent; Evans, Geraint; Robb, Nicole; Hwang, Ling Chin et al. (2016): RNA Polymerase Pausing during Initial Transcription. In *Molecular cell* 63 (6), pp. 939–950. DOI: 10.1016/j.molcel.2016.08.011.

Ebright, R. H. (2000): RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. In *Journal of molecular biology* 304 (5), pp. 687–698. DOI: 10.1006/jmbi.2000.4309.

Egly, Jean-Marc; Coin, Frederic (2011): A history of TFIIH: two decades of molecular biology on a pivotal transcription/repair factor. In *DNA repair* 10 (7), pp. 714–721. DOI: 10.1016/j.dnarep.2011.04.021.

Eichner, Jesse; Chen, Hung-Ta; Warfield, Linda; Hahn, Steven (2010): Position of the general transcription factor TFIIF within the RNA polymerase II transcription preinitiation complex. In *The EMBO journal* 29 (4), pp. 706–716. DOI: 10.1038/emboj.2009.386.

Engel, Christoph; Plitzko, Jurgen; Cramer, Patrick (2016): RNA polymerase I-Rrn3 complex at 4.8 A resolution. In *Nature communications* 7, p. 12129. DOI: 10.1038/ncomms12129.

Engel, Christoph; Sainsbury, Sarah; Cheung, Alan C.; Kostrewa, Dirk; Cramer, Patrick (2013): RNA polymerase I structure and transcription regulation. In *Nature* 502 (7473), pp. 650–655. DOI: 10.1038/nature12712.

Erie, Dorothy A. (2002): The many conformational states of RNA polymerase elongation complexes and their roles in the regulation of transcription. In *Biochimica et biophysica acta* 1577 (2), pp. 224–239.

Esyunina, Daria; Agapov, Aleksei; Kulbachinskiy, Andrey (2016): Regulation of transcriptional pausing through the secondary channel of RNA polymerase. In *Proceedings of the National Academy of Sciences of the United States of America* 113 (31), pp. 8699–8704. DOI: 10.1073/pnas.1603531113.

Facciotti, Marc T.; Reiss, David J.; Pan, Min; Kaur, Amardeep; Vuthoori, Madhavi; Bonneau, Richard et al. (2007): General transcription factor specified global gene regulation in archaea. In *Proceedings of the National Academy of Sciences of the United States of America* 104 (11), pp. 4630–4635. DOI: 10.1073/pnas.0611663104.

Faitar, S. L.; Brodie, S. A.; Ponticelli, A. S. (2001): Promoter-specific shifts in transcription initiation conferred by yeast TFIIB mutations are determined by the sequence in the immediate vicinity of the start sites. In *Molecular and cellular biology* 21 (14), pp. 4427–4440. DOI: 10.1128/MCB.21.14.4427-4440.2001.

Fang, S. M.; Burton, Z. F. (1996): RNA polymerase II-associated protein (RAP) 74 binds transcription factor (TF) IIB and blocks TFIIB-RAP30 binding. In *The Journal of biological chemistry* 271 (20), pp. 11703–11709.

Fanning, Ellen; Klimovich, Vitaly; Nager, Andrew R. (2006): A dynamic model for replication protein A (RPA) function in DNA processing pathways. In *Nucleic acids research* 34 (15), pp. 4126–4137. DOI: 10.1093/nar/gkl550.

Fatica, Alessandro; Bozzoni, Irene (2014): Long non-coding RNAs: new players in cell differentiation and development. In *Nature reviews. Genetics* 15 (1), pp. 7–21. DOI: 10.1038/nrg3606.

Feklistov, Andrey; Sharon, Brian D.; Darst, Seth A.; Gross, Carol A. (2014): Bacterial sigma factors: a historical, structural, and genomic perspective. In *Annual review of microbiology* 68, pp. 357–376. DOI: 10.1146/annurev-micro-092412-155737.

Feng, Yu; Zhang, Yu; Ebright, Richard H. (2016): Structural basis of transcription activation. In *Science (New York, N.Y.)* 352 (6291), pp. 1330–1333. DOI: 10.1126/science.aaf4417.

Fiala, Gerhard; Stetter, Karl O. (1986): Pyrococcus furiosus sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100°C. In *Arch. Microbiol.* 145 (1), pp. 56–61. DOI: 10.1007/BF00413027.

Fishburn, James; Tomko, Eric; Galburt, Eric; Hahn, Steven (2015): Double-stranded DNA translocase activity of transcription factor TFIIH and the mechanism of RNA polymerase II open complex formation. In *Proceedings of the National Academy of Sciences of the United States of America* 112 (13), pp. 3961–3966. DOI: 10.1073/pnas.1417709112.

Flores, O.; Lu, H.; Killeen, M.; Greenblatt, J.; Burton, Z. F.; Reinberg, D. (1991): The small subunit of transcription factor IIF recruits RNA polymerase II into the preinitiation complex. In *Proceedings of the National Academy of Sciences of the United States of America* 88 (22), pp. 9999–10003.

Flores, O.; Maldonado, E.; Burton, Z.; Greenblatt, J.; Reinberg, D. (1988): Factors involved in specific transcription by mammalian RNA polymerase II. RNA polymerase II-associating protein 30 is an essential component of transcription factor IIF. In *The Journal of biological chemistry* 263 (22), pp. 10812–10816.

Forget, Diane; Langelier, Marie-France; Therien, Cynthia; Trinh, Vincent; Coulombe, Benoit (2004): Photo-cross-linking of a purified preinitiation complex reveals central roles for the RNA polymerase II mobile clamp and TFIIE in initiation mechanisms. In *Molecular and cellular biology* 24 (3), pp. 1122–1131.

Forterre, P.; Bergerat, A.; Lopez-Garcia, P. (1996): The unique DNA topology and DNA topoisomerases of hyperthermophilic archaea. In *FEMS microbiology reviews* 18 (2-3), pp. 237–248.

Fouqueau, Thomas; Zeller, Mirijam E.; Cheung, Alan C.; Cramer, Patrick; Thomm, Michael (2013): The RNA polymerase trigger loop functions in all three phases of the transcription cycle. In *Nucleic acids research* 41 (14), pp. 7048–7059. DOI: 10.1093/nar/gkt433.

Franklin, C. C.; McCulloch, A. V.; Kraft, A. S. (1995): In vitro association between the Jun protein family and the general transcription factors, TBP and TFIIB. In *The Biochemical journal* 305 (Pt 3), pp. 967–974.

Gaiser, F.; Tan, S.; Richmond, T. J. (2000): Novel dimerization fold of RAP30/RAP74 in human TFIIF at 1.7 A resolution. In *Journal of molecular biology* 302 (5), pp. 1119–1127. DOI: 10.1006/jmbi.2000.4110.

Garcia-Lopez, M. Carmen; Navarro, Francisco (2011): RNA polymerase II conserved protein domains as platforms for protein-protein interactions. In *Transcription* 2 (4), pp. 193–197. DOI: 10.4161/trns.2.4.16786.

Gehring, Alexandra M.; Walker, Julie E.; Santangelo, Thomas J. (2016): Transcription Regulation in Archaea. In *Journal of bacteriology* 198 (14), pp. 1906–1917. DOI: 10.1128/JB.00255-16.

Geiger, J. H.; Hahn, S.; Lee, S.; Sigler, P. B. (1996): Crystal structure of the yeast TFIIA/TBP/DNA complex. In *Science (New York, N.Y.)* 272 (5263), pp. 830–836.

Geiger, Sebastian R.; Lorenzen, Kristina; Schreieck, Amelie; Hanecker, Patrizia; Kostrewa, Dirk; Heck, Albert J. R.; Cramer, Patrick (2010): RNA polymerase I contains a TFIIF-related DNA-binding subcomplex. In *Molecular cell* 39 (4), pp. 583–594. DOI: 10.1016/j.molcel.2010.07.028.

Ghazy, Mohamed A.; Brodie, Seth A.; Ammerman, Michelle L.; Ziegler, Lynn M.; Ponticelli, Alfred S. (2004): Amino acid substitutions in yeast TFIIF confer upstream shifts in transcription initiation and altered interaction with RNA polymerase II. In *Molecular and cellular biology* 24 (24), pp. 10975–10985. DOI: 10.1128/MCB.24.24.10975-10985.2004.

Ghildiyal, Megha; Zamore, Phillip D. (2009): Small silencing RNAs: an expanding universe. In *Nature reviews. Genetics* 10 (2), pp. 94–108. DOI: 10.1038/nrg2504.

Giardina, C.; Lis, J. T. (1993): DNA melting on yeast RNA polymerase II promoters. In *Science (New York, N.Y.)* 261 (5122), pp. 759–762.

Gibbons, Brian J.; Brignole, Edward J.; Azubel, Maia; Murakami, Kenji; Voss, Neil R.; Bushnell, David A. et al. (2012): Subunit architecture of general transcription factor TFIIH. In *Proceedings of the National Academy of Sciences of the United States of America* 109 (6), pp. 1949–1954. DOI: 10.1073/pnas.1105266109.

Gietl, Andreas; Holzmeister, Phil; Blombach, Fabian; Schulz, Sarah; Voithenberg, Lena Voith von; Lamb, Don C. et al. (2014): Eukaryotic and archaeal TBP and TFB/TF(II)B follow different promoter DNA bending pathways. In *Nucleic acids research* 42 (10), pp. 6219–6231. DOI: 10.1093/nar/gku273.

Gindner, Antonia; Hausner, Winfried; Thomm, Michael (2014): The TrmB family: a versatile group of transcriptional regulators in Archaea. In *Extremophiles : life under extreme conditions* 18 (5), pp. 925–936. DOI: 10.1007/s00792-014-0677-2.

Gnatt, A. L.; Cramer, P.; Fu, J.; Bushnell, D. A.; Kornberg, R. D. (2001): Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. In *Science (New York, N.Y.)* 292 (5523), pp. 1876–1882. DOI: 10.1126/science.1059495.

Goede, Bernd (2004): Protein-Protein-Wechselwirkungen der Untereinheiten der RNA-Polymerase von Pyrococcus furiosus. Available online at http://macau.uni-kiel.de/receive/dissertation_diss_00001046.

Goodrich, J. A.; Tjian, R. (1994): Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. In *Cell* 77 (1), pp. 145–156.

Gorecki, Andrzej; Figiel, Malgorzata; Dziedzicka-Wasylewska, Marta (2015): In vitro fluorescence studies of transcription factor IIB-DNA interaction. In *Acta biochimica Polonica* 62 (3), pp. 413–421. DOI: 10.18388/abp.2015_1034.

Grohmann, Dina; Hirtreiter, Angela; Werner, Finn (2009): RNAP subunits F/E (RPB4/7) are stably associated with archaeal RNA polymerase: using fluorescence anisotropy to monitor RNAP assembly in vitro. In *The Biochemical journal* 421 (3), pp. 339–343. DOI: 10.1042/BJ20090782.

Grohmann, Dina; Nagy, Julia; Chakraborty, Anirban; Klose, Daniel; Fielden, Daniel; Ebright, Richard H. et al. (2011): The initiation factor TFE and the elongation factor Spt4/5 compete for the RNAP clamp

during transcription initiation and elongation. In *Molecular cell* 43 (2), pp. 263–274. DOI: 10.1016/j.molcel.2011.05.030.

Grohmann, Dina; Werner, Finn (2011): Recent advances in the understanding of archaeal transcription. In *Current opinion in microbiology* 14 (3), pp. 328–334. DOI: 10.1016/j.mib.2011.04.012.

Grossman, A. D.; Erickson, J. W.; Gross, C. A. (1984): The htpR gene product of E. coli is a sigma factor for heat-shock promoters. In *Cell* 38 (2), pp. 383–390.

Grünberg, Sebastian (2009): Untersuchungen zur Funktion der RNA Polymerase Untereinheit E' und des Transkriptionsfaktors E. Available online at http://epub.uni-regensburg.de/12103/.

Grünberg, Sebastian; Bartlett, Michael S.; Naji, Souad; Thomm, Michael (2007): Transcription factor E is a part of transcription elongation complexes. In *The Journal of biological chemistry* 282 (49), pp. 35482–35490. DOI: 10.1074/jbc.M707371200.

Grünberg, Sebastian; Reich, Christoph; Zeller, Mirijam E.; Bartlett, Michael S.; Thomm, Michael (2010): Rearrangement of the RNA polymerase subunit H and the lower jaw in archaeal elongation complexes. In *Nucleic acids research* 38 (6), pp. 1950–1963. DOI: 10.1093/nar/gkp1190.

Grünberg, Sebastian; Warfield, Linda; Hahn, Steven (2012): Architecture of the RNA polymerase II preinitiation complex and mechanism of ATP-dependent promoter opening. In *Nature structural & molecular biology* 19 (8), pp. 788–796. DOI: 10.1038/nsmb.2334.

Guo, Li; Feng, Yingang; Zhang, ZhenFeng; Yao, Hongwei; Luo, Yuanming; Wang, Jinfeng; Huang, Li (2008): Biochemical and structural characterization of Cren7, a novel chromatin protein conserved among Crenarchaea. In *Nucleic acids research* 36 (4), pp. 1129–1137. DOI: 10.1093/nar/gkm1128.

Ha, I.; Lane, W. S.; Reinberg, D. (1991): Cloning of a human gene encoding the general transcription initiation factor IIB. In *Nature* 352 (6337), pp. 689–695. DOI: 10.1038/352689a0.

Ha, I.; Roberts, S.; Maldonado, E.; Sun, X.; Kim, L. U.; Green, M.; Reinberg, D. (1993): Multiple functional domains of human transcription factor IIB: distinct interactions with two general transcription factors and RNA polymerase II. In *Genes & development* 7 (6), pp. 1021–1032.

Hadzic, E.; Desai-Yajnik, V.; Helmer, E.; Guo, S.; Wu, S.; Koudinova, N. et al. (1995): A 10-amino-acid sequence in the N-terminal A/B domain of thyroid hormone receptor alpha is essential for transcriptional activation and interaction with the general transcription factor TFIIB. In *Molecular and cellular biology* 15 (8), pp. 4507–4517.

Hampsey, M. (1998): Molecular genetics of the RNA polymerase II general transcriptional machinery. In *Microbiology and molecular biology reviews : MMBR* 62 (2), pp. 465–503.

Hanzelka, B. L.; Darcy, T. J.; Reeve, J. N. (2001): TFE, an archaeal transcription factor in Methanobacterium thermoautotrophicum related to eucaryal transcription factor TFIIEalpha. In *Journal of bacteriology* 183 (5), pp. 1813–1818. DOI: 10.1128/JB.183.5.1813-1818.2001.

Hausner, W.; Frey, G.; Thomm, M. (1991): Control regions of an archaeal gene. A TATA box and an initiator element promote cell-free transcription of the tRNA(Val) gene of Methanococcus vannielii. In *Journal of molecular biology* 222 (3), pp. 495–508.

Hausner, W.; Lange, U.; Musfeldt, M. (2000): Transcription factor S, a cleavage induction factor of the archaeal RNA polymerase. In *The Journal of biological chemistry* 275 (17), pp. 12393–12399.

Hausner, W.; Thomm, M. (2001): Events during initiation of archaeal transcription: open complex formation and DNA-protein interactions. In *Journal of bacteriology* 183 (10), pp. 3025–3031. DOI: 10.1128/JB.183.10.3025-3031.2001.

Hausner, W.; Wettach, J.; Hethke, C.; Thomm, M. (1996): Two transcription factors related with the eucaryal transcription factors TATA-binding protein and transcription factor IIB direct promoter recognition by an archaeal RNA polymerase. In *The Journal of biological chemistry* 271 (47), pp. 30144–30148.

Hayashi, Kazuhiro; Watanabe, Tomomichi; Tanaka, Aki; Furumoto, Tadashi; Sato-Tsuchiya, Chiaki; Kimura, Makoto et al. (2005): Studies of Schizosaccharomyces pombe TFIIE indicate conformational and functional changes in RNA polymerase II at transcription initiation. In *Genes to cells : devoted to molecular & cellular mechanisms* 10 (3), pp. 207–224. DOI: 10.1111/j.1365-2443.2005.00833.x.

He, Yuan; Fang, Jie; Taatjes, Dylan J.; Nogales, Eva (2013): Structural visualization of key steps in human transcription initiation. In *Nature* 495 (7442), pp. 481–486. DOI: 10.1038/nature11991.

HERSHEY, A. D.; CHASE, M. (1952): Independent functions of viral protein and nucleic acid in growth of bacteriophage. In *The Journal of general physiology* 36 (1), pp. 39–56.

Hethke, C.; Geerling, A. C.; Hausner, W.; Vos, W. M. de; Thomm, M. (1996): A cell-free transcription system for the hyperthermophilic archaeon Pyrococcus furiosus. In *Nucleic acids research* 24 (12), pp. 2369–2376.

Hirata, Akira; Klein, Brianna J.; Murakami, Katsuhiko S. (2008): The X-ray crystal structure of RNA polymerase from Archaea. In *Nature* 451 (7180), pp. 851–854. DOI: 10.1038/nature06530.

Hirtreiter, Angela; Damsma, Gerke E.; Cheung, Alan C. M.; Klose, Daniel; Grohmann, Dina; Vojnic, Erika et al. (2010a): Spt4/5 stimulates transcription elongation through the RNA polymerase clamp coiled-coil motif. In *Nucleic acids research* 38 (12), pp. 4040–4051. DOI: 10.1093/nar/gkq135.

Hirtreiter, Angela; Grohmann, Dina; Werner, Finn (2010b): Molecular mechanisms of RNA polymerase--the F/E (RPB4/7) complex is required for high processivity in vitro. In *Nucleic acids research* 38 (2), pp. 585–596. DOI: 10.1093/nar/gkp928.

Hoffmann, Niklas A.; Sadian, Yashar; Tafur, Lucas; Kosinski, Jan; Muller, Christoph W. (2016): Specialization versus conservation: How Pol I and Pol III use the conserved architecture of the pre-initiation complex for specialized transcription. In *Transcription* 7 (4), pp. 127–132. DOI: 10.1080/21541264.2016.1203628.

Holmberg, Anders; Blomstergren, Anna; Nord, Olof; Lukacs, Morten; Lundeberg, Joakim; Uhlen, Mathias (2005): The biotin-streptavidin interaction can be reversibly broken using water at elevated temperatures. In *Electrophoresis* 26 (3), pp. 501–510. DOI: 10.1002/elps.200410070.

Holstege, F. C.; van der Vliet, P C; Timmers, H. T. (1996): Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH. In *The EMBO journal* 15 (7), pp. 1666–1677.

Hsin, Jing-Ping; Manley, James L. (2012): The RNA polymerase II CTD coordinates transcription and RNA processing. In *Genes & development* 26 (19), pp. 2119–2137. DOI: 10.1101/gad.200303.112.

Hsu, Lilian M. (2002): Promoter clearance and escape in prokaryotes. In *Biochimica et biophysica acta* 1577 (2), pp. 191–207.

Imbalzano, A. N.; Zaret, K. S.; Kingston, R. E. (1994): Transcription factor (TF) IIB and TFIIA can independently increase the affinity of the TATA-binding protein for DNA. In *The Journal of biological chemistry* 269 (11), pp. 8280–8286.

Izban, M. G.; Luse, D. S. (1992): The RNA polymerase II ternary complex cleaves the nascent transcript in a 3'----5' direction in the presence of elongation factor SII. In *Genes & development* 6 (7), pp. 1342–1356.

Jelinska, Clare; Conroy, Matthew J.; Craven, C. Jeremy; Hounslow, Andrea M.; Bullough, Per A.; Waltho, Jonathan P. et al. (2005): Obligate heterodimerization of the archaeal Alba2 protein with Alba1 provides a mechanism for control of DNA packaging. In *Structure (London, England : 1993)* 13 (7), pp. 963–971. DOI: 10.1016/j.str.2005.04.016.

Jeronimo, Celia; Langelier, Marie-France; Zeghouf, Mahel; Cojocaru, Marilena; Bergeron, Dominique; Baali, Dania et al. (2004): RPAP1, a novel human RNA polymerase II-associated protein affinity purified with recombinant wild-type and mutated polymerase subunits. In *Molecular and cellular biology* 24 (16), pp. 7043–7058. DOI: 10.1128/MCB.24.16.7043-7058.2004.

Jun, Sung-Hoon; Hirata, Akira; Kanai, Tamotsu; Santangelo, Thomas J.; Imanaka, Tadayuki; Murakami, Katsuhiko S. (2014): The X-ray crystal structure of the euryarchaeal RNA polymerase in an open-clamp configuration. In *Nature communications* 5, p. 5132. DOI: 10.1038/ncomms6132.

Juo, Z. S.; Chiu, T. K.; Leiberman, P. M.; Baikalov, I.; Berk, A. J.; Dickerson, R. E. (1996): How proteins recognize the TATA box. In *Journal of molecular biology* 261 (2), pp. 239–254. DOI: 10.1006/jmbi.1996.0456.

Juven-Gershon, Tamar; Cheng, Susan; Kadonaga, James T. (2006): Rational design of a super core promoter that enhances gene expression. In *Nature methods* 3 (11), pp. 917–922.

Juven-Gershon, Tamar; Hsu, Jer-Yuan; Theisen, Joshua Wm; Kadonaga, James T. (2008): The RNA polymerase II core promoter - the gateway to transcription. In *Current opinion in cell biology* 20 (3), pp. 253–259. DOI: 10.1016/j.ceb.2008.03.003.

Kamada, K.; Angelis, J. de; Roeder, R. G.; Burley, S. K. (2001): Crystal structure of the C-terminal domain of the RAP74 subunit of human transcription factor IIF. In *Proceedings of the National Academy of Sciences of the United States of America* 98 (6), pp. 3115–3120. DOI: 10.1073/pnas.051631098.

Kang, J. J.; Auble, D. T.; Ranish, J. A.; Hahn, S. (1995): Analysis of the yeast transcription factor TFIIA: distinct functional regions and a polymerase II-specific role in basal and activated transcription. In *Molecular and cellular biology* 15 (3), pp. 1234–1243.

Kassavetis, G. A.; Braun, B. R.; Nguyen, L. H.; Geiduschek, E. P. (1990): S. cerevisiae TFIIIB is the transcription initiation factor proper of RNA polymerase III, while TFIIIA and TFIIIC are assembly factors. In *Cell* 60 (2), pp. 235–245.

Kassavetis, G. A.; Geiduschek, E. P. (2006): Transcription factor TFIIIB and transcription by RNA polymerase III. In *Biochemical Society transactions* 34 (Pt 6), pp. 1082–1087. DOI: 10.1042/BST0341082.

Kassavetis, G. A.; Letts, G. A.; Geiduschek, E. P. (2001): The RNA polymerase III transcription initiation factor TFIIIB participates in two steps of promoter opening. In *The EMBO journal* 20 (11), pp. 2823–2834. DOI: 10.1093/emboj/20.11.2823.

Kauer, J. C.; Erickson-Viitanen, S.; Wolfe, H. R., JR; DeGrado, W. F. (1986): p-Benzoyl-L-phenylalanine, a new photoreactive amino acid. Photolabeling of calmodulin with a synthetic calmodulin-binding peptide. In *The Journal of biological chemistry* 261 (23), pp. 10695–10700.

Kerr, Iain D.; Wadsworth, Ross I. M.; Cubeddu, Liza; Blankenfeldt, Wulf; Naismith, James H.; White, Malcolm F. (2003): Insights into ssDNA recognition by the OB fold from a structural and thermodynamic study of Sulfolobus SSB protein. In *The EMBO journal* 22 (11), pp. 2561–2570. DOI: 10.1093/emboj/cdg272.

Kettenberger, Hubert; Armache, Karim-Jean; Cramer, Patrick (2004): Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. In *Molecular cell* 16 (6), pp. 955–965. DOI: 10.1016/j.molcel.2004.11.040.

Khaperskyy, Denys A.; Ammerman, Michelle L.; Majovski, Robert C.; Ponticelli, Alfred S. (2008): Functions of Saccharomyces cerevisiae TFIIF during transcription start site utilization. In *Molecular and cellular biology* 28 (11), pp. 3757–3766. DOI: 10.1128/MCB.02272-07.

Kim, J. L.; Nikolov, D. B.; Burley, S. K. (1993a): Co-crystal structure of TBP recognizing the minor groove of a TATA element. In *Nature* 365 (6446), pp. 520–527. DOI: 10.1038/365520a0.

Kim, T. K.; Ebright, R. H.; Reinberg, D. (2000): Mechanism of ATP-dependent promoter melting by transcription factor IIH. In *Science (New York, N.Y.)* 288 (5470), pp. 1418–1422.

Kim, Y.; Geiger, J. H.; Hahn, S.; Sigler, P. B. (1993b): Crystal structure of a yeast TBP/TATA-box complex. In *Nature* 365 (6446), pp. 512–520. DOI: 10.1038/365512a0.

Kireeva, Maria L.; Domecq, Celine; Coulombe, Benoit; Burton, Zachary F.; Kashlev, Mikhail (2011): Interaction of RNA polymerase II fork loop 2 with downstream non-template DNA regulates transcription elongation. In *The Journal of biological chemistry* 286 (35), pp. 30898–30910. DOI: 10.1074/jbc.M111.260844.

Komori, K.; Ishino, Y. (2001): Replication protein A in Pyrococcus furiosus is involved in homologous DNA recombination. In *The Journal of biological chemistry* 276 (28), pp. 25654–25660. DOI: 10.1074/jbc.M102423200.

Koonin, Eugene V. (2015): Why the Central Dogma: on the nature of the great biological exclusion principle. In *Biology direct* 10, p. 52. DOI: 10.1186/s13062-015-0084-3.

Korkhin, Yakov; Unligil, Ulug M.; Littlefield, Otis; Nelson, Pamlea J.; Stuart, David I.; Sigler, Paul B. et al. (2009): Evolution of complex RNA polymerases: the complete archaeal RNA polymerase structure. In *PLoS biology* 7 (5), e1000102. DOI: 10.1371/journal.pbio.1000102.

Kornberg, Roger D. (2007): The molecular basis of eukaryotic transcription. In *Proceedings of the National Academy of Sciences of the United States of America* 104 (32), pp. 12955–12961. DOI: 10.1073/pnas.0704138104.

Kosa, P. F.; Ghosh, G.; DeDecker, B. S.; Sigler, P. B. (1997): The 2.1-A crystal structure of an archaeal preinitiation complex: TATA-box-binding protein/transcription factor (II)B core/TATA-box. In *Proceedings of the National Academy of Sciences of the United States of America* 94 (12), pp. 6042–6047.

Kostrewa, Dirk; Zeller, Mirijam E.; Armache, Karim-Jean; Seizl, Martin; Leike, Kristin; Thomm, Michael; Cramer, Patrick (2009): RNA polymerase II-TFIIB structure and mechanism of transcription initiation. In *Nature* 462 (7271), pp. 323–330. DOI: 10.1038/nature08548.

Krogan, Nevan J.; Cagney, Gerard; Yu, Haiyuan; Zhong, Gouqing; Guo, Xinghua; Ignatchenko, Alexandr et al. (2006): Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. In *Nature* 440 (7084), pp. 637–643. DOI: 10.1038/nature04670.

Kuehner, Jason N.; Brow, David A. (2006): Quantitative analysis of in vivo initiator selection by yeast RNA polymerase II supports a scanning model. In *The Journal of biological chemistry* 281 (20), pp. 14119–14128. DOI: 10.1074/jbc.M601937200.

Kuhn, Claus-D; Geiger, Sebastian R.; Baumli, Sonja; Gartmann, Marco; Gerber, Jochen; Jennebach, Stefan et al. (2007): Functional architecture of RNA polymerase I. In *Cell* 131 (7), pp. 1260–1272. DOI: 10.1016/j.cell.2007.10.051.

Kuldell, N. H.; Buratowski, S. (1997): Genetic analysis of the large subunit of yeast transcription factor IIE reveals two regions with distinct functions. In *Molecular and cellular biology* 17 (9), pp. 5288–5298.

Kuznedelov, Konstantin; Korzheva, Nataliya; Mustaev, Arkady; Severinov, Konstantin (2002): Structure-based analysis of RNA polymerase function: the largest subunit's rudder contributes critically to elongation complex stability and is not involved in the maintenance of RNA-DNA hybrid length. In *The EMBO journal* 21 (6), pp. 1369–1378. DOI: 10.1093/emboj/21.6.1369.

Kyrpides, N. C.; Ouzounis, C. A. (1999): Transcription in archaea. In *Proceedings of the National Academy of Sciences of the United States of America* 96 (15), pp. 8545–8550.

Laemmli, U. K. (1970): Cleavage of structural proteins during the assembly of the head of bacteriophage T4. In *Nature* 227 (5259), pp. 680–685.

Lagrange, T.; Kapanidis, A. N.; Tang, H.; Reinberg, D.; Ebright, R. H. (1998): New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. In *Genes & development* 12 (1), pp. 34–44.

Landick, Robert (2009): Functional divergence in the growing family of RNA polymerases. In *Structure (London, England : 1993)* 17 (3), pp. 323–325. DOI: 10.1016/j.str.2009.02.006.

Landrieux, Emilie; Alic, Nazif; Ducrot, Cecile; Acker, Joel; Riva, Michel; Carles, Christophe (2006): A subcomplex of RNA polymerase III subunits involved in transcription termination and reinitiation. In *The EMBO journal* 25 (1), pp. 118–128. DOI: 10.1038/sj.emboj.7600915.

Laurens, Niels; Driessen, Rosalie P. C.; Heller, Iddo; Vorselen, Daan; Noom, Maarten C.; Hol, Felix J. H. et al. (2012): Alba shapes the archaeal genome using a delicate balance of bridging and stiffening the DNA. In *Nature communications* 3, p. 1328. DOI: 10.1038/ncomms2330.

Lee, Dong-Hoon; Gershenzon, Naum; Gupta, Malavika; Ioshikhes, Ilya P.; Reinberg, Danny; Lewis, Brian A. (2005): Functional characterization of core promoter elements: the downstream core element is recognized by TAF1. In *Molecular and cellular biology* 25 (21), pp. 9674–9686. DOI: 10.1128/MCB.25.21.9674-9686.2005.

Lee, S.; Hahn, S. (1995): Model for binding of transcription factor TFIIB to the TBP-DNA complex. In *Nature* 376 (6541), pp. 609–612. DOI: 10.1038/376609a0.

Leurent, Claire; Sanders, Steven L.; Demeny, Mate A.; Garbett, Krassimira A.; Ruhlmann, Christine; Weil, P. Anthony et al. (2004): Mapping key functional sites within yeast TFIID. In *The EMBO journal* 23 (4), pp. 719–727. DOI: 10.1038/sj.emboj.7600111.

Lewis, B. A.; Kim, T. K.; Orkin, S. H. (2000): A downstream element in the human beta-globin promoter: evidence of extended sequence-specific transcription factor IID contacts. In *Proceedings of the National Academy of Sciences of the United States of America* 97 (13), pp. 7172–7177. DOI: 10.1073/pnas.120181197.

Li, Y.; Flanagan, P. M.; Tschochner, H.; Kornberg, R. D. (1994): RNA polymerase II initiation factor interactions and transcription start site selection. In *Science (New York, N.Y.)* 263 (5148), pp. 805–807.

Lim, Chin Yan; Santoso, Buyung; Boulay, Thomas; Dong, Emily; Ohler, Uwe; Kadonaga, James T. (2004): The MTE, a new core promoter element for transcription by RNA polymerase II. In *Genes & development* 18 (13), pp. 1606–1617. DOI: 10.1101/gad.1193404.

Lin, Yin Chun; Choi, Wai S.; Gralla, Jay D. (2005): TFIIH XPB mutants suggest a unified bacterial-like mechanism for promoter opening but not escape. In *Nature structural & molecular biology* 12 (7), pp. 603–607. DOI: 10.1038/nsmb949.

Lisica, Ana; Engel, Christoph; Jahnel, Marcus; Roldan, Edgar; Galburt, Eric A.; Cramer, Patrick; Grill, Stephan W. (2016): Mechanisms of backtrack recovery by RNA polymerases I and II. In *Proceedings of the National Academy of Sciences of the United States of America* 113 (11), pp. 2946–2951. DOI: 10.1073/pnas.1517011113.

Liu, Xin; Bushnell, David A.; Wang, Dong; Calero, Guillermo; Kornberg, Roger D. (2010): Structure of an RNA polymerase II-TFIIB complex and the transcription initiation mechanism. In *Science (New York, N.Y.)* 327 (5962), pp. 206–209. DOI: 10.1126/science.1182015.

Lonetto, M.; Gribskov, M.; Gross, C. A. (1992): The sigma 70 family: sequence conservation and evolutionary relationships. In *Journal of bacteriology* 174 (12), pp. 3843–3849.

Lozinski, Tomasz; Wierzchowski, Kazimierz L. (2003): Inactivation and destruction by KMnO4 of Escherichia coli RNA polymerase open transcription complex: recommendations for footprinting experiments. In *Analytical biochemistry* 320 (2), pp. 239–251.

Luger, K.; Mader, A. W.; Richmond, R. K.; Sargent, D. F.; Richmond, T. J. (1997): Crystal structure of the nucleosome core particle at 2.8 A resolution. In *Nature* 389 (6648), pp. 251–260. DOI: 10.1038/38444.

Luse, Donal S. (2013): Promoter clearance by RNA polymerase II. In *Biochimica et biophysica acta* 1829 (1), pp. 63–68. DOI: 10.1016/j.bbagrm.2012.08.010.

MacDonald, P. N.; Sherman, D. R.; Dowd, D. R.; Jefcoat, S. C., JR; DeLisle, R. K. (1995): The vitamin D receptor interacts with general transcription factor IIB. In *The Journal of biological chemistry* 270 (9), pp. 4748–4752.

Malik, S.; Hisatake, K.; Sumimoto, H.; Horikoshi, M.; Roeder, R. G. (1991): Sequence of general transcription factor TFIIB and relationships to other initiation factors. In *Proceedings of the National Academy of Sciences of the United States of America* 88 (21), pp. 9553–9557.

Malik, S.; Lee, D. K.; Roeder, R. G. (1993): Potential RNA polymerase II-induced interactions of transcription factor TFIIB. In *Molecular and cellular biology* 13 (10), pp. 6253–6259.

Marsh, T. L.; Reich, C. I.; Whitelock, R. B.; Olsen, G. J. (1994): Transcription factor IID in the Archaea: sequences in the Thermococcus celer genome would encode a product closely related to the TATA-binding protein of eukaryotes. In *Proceedings of the National Academy of Sciences of the United States of America* 91 (10), pp. 4180–4184.

Maruyama, Hugo; Shin, Minsang; Oda, Toshiyuki; Matsumi, Rie; Ohniwa, Ryosuke L.; Itoh, Takehiko et al. (2011): Histone and TK0471/TrmBL2 form a novel heterogeneous genome architecture in the hyperthermophilic archaeon Thermococcus kodakarensis. In *Molecular biology of the cell* 22 (3), pp. 386–398. DOI: 10.1091/mbc.E10-08-0668.

Masters, Kristina M.; Parkhurst, Kay M.; Daugherty, Margaret A.; Parkhurst, Lawrence J. (2003): Native human TATA-binding protein simultaneously binds and bends promoter DNA without a slow isomerization step or TFIIB requirement. In *The Journal of biological chemistry* 278 (34), pp. 31685–31690. DOI: 10.1074/jbc.M305201200.

Matangkasombut, Oranart; Auty, Roy; Buratowski, Stephen (2004): Structure and function of the TFIID complex. In *Advances in protein chemistry* 67, pp. 67–92. DOI: 10.1016/S0065-3233(04)67003-3.

Maxon, M. E.; Goodrich, J. A.; Tjian, R. (1994): Transcription factor IIE binds preferentially to RNA polymerase IIa and recruits TFIIH: a model for promoter clearance. In *Genes & development* 8 (5), pp. 515–524.

Micorescu, Michael; Grunberg, Sebastian; Franke, Andreas; Cramer, Patrick; Thomm, Michael; Bartlett, Michael (2008): Archaeal transcription: function of an alternative transcription factor B from Pyrococcus furiosus. In *Journal of bacteriology* 190 (1), pp. 157–167. DOI: 10.1128/JB.01498-07.

Miller, G.; Panov, K. I.; Friedrich, J. K.; Trinkle-Mulcahy, L.; Lamond, A. I.; Zomerdijk, J. C. (2001): hRRN3 is essential in the SL1-mediated recruitment of RNA Polymerase I to rRNA gene promoters. In *The EMBO journal* 20 (6), pp. 1373–1382. DOI: 10.1093/emboj/20.6.1373.

Mooney, Rachel Anne; Schweimer, Kristian; Rosch, Paul; Gottesman, Max; Landick, Robert (2009): Two structurally independent domains of E. coli NusG create regulatory plasticity via distinct interactions with RNA polymerase and regulators. In *Journal of molecular biology* 391 (2), pp. 341–358. DOI: 10.1016/j.jmb.2009.05.078.

Müller, Ferenc; Demeny, Mate A.; Tora, Laszlo (2007): New problems in RNA polymerase II transcription initiation: matching the diversity of core promoters with a variety of promoter recognition factors. In *The Journal of biological chemistry* 282 (20), pp. 14685–14689. DOI: 10.1074/jbc.R700012200.

Murzin, A. G. (1993): OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. In *The EMBO journal* 12 (3), pp. 861–867.

Na, J. G.; Hampsey, M. (1993): The Kluyveromyces gene encoding the general transcription factor IIB: structural analysis and expression in Saccharomyces cerevisiae. In *Nucleic acids research* 21 (15), pp. 3413–3417.

Nagy, Julia; Grohmann, Dina; Cheung, Alan C. M.; Schulz, Sarah; Smollett, Katherine; Werner, Finn; Michaelis, Jens (2015): Complete architecture of the archaeal RNA polymerase open complex from single-molecule FRET and NPS. In *Nature communications* 6, p. 6161. DOI: 10.1038/ncomms7161.

Naidu, Srivatsava; Friedrich, J. Karsten; Russell, Jackie; Zomerdijk, Joost C B M (2011): TAF1B is a TFIIB-like component of the basal transcription machinery for RNA polymerase I. In *Science (New York, N.Y.)* 333 (6049), pp. 1640–1642. DOI: 10.1126/science.1207656.

Naji, Souad; Grunberg, Sebastian; Thomm, Michael (2007): The RPB7 orthologue E' is required for transcriptional activity of a reconstituted archaeal core enzyme at low temperatures and stimulates open complex formation. In *The Journal of biological chemistry* 282 (15), pp. 11047–11057. DOI: 10.1074/jbc.M611674200.

Nalabothula, Narasimharao; Xi, Liqun; Bhattacharyya, Sucharita; Widom, Jonathan; Wang, Ji-Ping; Reeve, John N. et al. (2013): Archaeal nucleosome positioning in vivo and in vitro is directed by primary sequence motifs. In *BMC genomics* 14, p. 391. DOI: 10.1186/1471-2164-14-391.

Naryshkina, Tatyana; Kuznedelov, Konstantin; Severinov, Konstantin (2006): The role of the largest RNA polymerase subunit lid element in preventing the formation of extended RNA-DNA hybrid. In *Journal of molecular biology* 361 (4), pp. 634–643. DOI: 10.1016/j.jmb.2006.05.034.

Nielsen, Soren; Yuzenkova, Yulia; Zenkin, Nikolay (2013): Mechanism of eukaryotic RNA polymerase III transcription termination. In *Science (New York, N.Y.)* 340 (6140), pp. 1577–1580. DOI: 10.1126/science.1237934.

Nikolov, D. B.; Chen, H.; Halay, E. D.; Usheva, A. A.; Hisatake, K.; Lee, D. K. et al. (1995): Crystal structure of a TFIIB-TBP-TATA-element ternary complex. In *Nature* 377 (6545), pp. 119–128. DOI: 10.1038/377119a0.

Nudler, E.; Mustaev, A.; Lukhtanov, E.; Goldfarb, A. (1997): The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. In *Cell* 89 (1), pp. 33–41.

Ohkuma, Y.; Hashimoto, S.; Wang, C. K.; Horikoshi, M.; Roeder, R. G. (1995): Analysis of the role of TFIIE in basal transcription and TFIIH-mediated carboxy-terminal domain phosphorylation through structure-function studies of TFIIE-alpha. In *Molecular and cellular biology* 15 (9), pp. 4856–4866.

Orioli, Andrea; Pascali, Chiara; Quartararo, Jade; Diebel, Kevin W.; Praz, Viviane; Romascano, David et al. (2011): Widespread occurrence of non-canonical transcription termination by human RNA polymerase III. In *Nucleic acids research* 39 (13), pp. 5499–5512. DOI: 10.1093/nar/gkr074.

Orlicky, S. M.; Tran, P. T.; Sayre, M. H.; Edwards, A. M. (2001): Dissociable Rpb4-Rpb7 subassembly of rna polymerase II binds to single-strand nucleic acid and mediates a post-recruitment step in transcription initiation. In *The Journal of biological chemistry* 276 (13), pp. 10097–10102. DOI: 10.1074/jbc.M003165200.

Orphanides, G.; Lagrange, T.; Reinberg, D. (1996): The general transcription factors of RNA polymerase II. In *Genes & development* 10 (21), pp. 2657–2683.

Ouhammouch, Mohamed; Dewhurst, Robert E.; Hausner, Winfried; Thomm, Michael; Geiduschek, E. Peter (2003): Activation of archaeal transcription by recruitment of the TATA-binding protein. In *Proceedings of the National Academy of Sciences of the United States of America* 100 (9), pp. 5097–5102. DOI: 10.1073/pnas.0837150100.

Paget, Mark S. (2015): Bacterial Sigma Factors and Anti-Sigma Factors: Structure, Function and Distribution. In *Biomolecules* 5 (3), pp. 1245–1265. DOI: 10.3390/biom5031245.

Pal, Mahadeb; Luse, Donal S. (2003): The initiation-elongation transition: lateral mobility of RNA in RNA polymerase II complexes is greatly reduced at +8/+9 and absent by +23. In *Proceedings of the National Academy of Sciences of the United States of America* 100 (10), pp. 5700–5705. DOI: 10.1073/pnas.1037057100.

Pal, Mahadeb; Ponticelli, Alfred S.; Luse, Donal S. (2005): The role of the transcription bubble and TFIIB in promoter clearance by RNA polymerase II. In *Molecular cell* 19 (1), pp. 101–110. DOI: 10.1016/j.molcel.2005.05.024.

Pan, G.; Greenblatt, J. (1994): Initiation of transcription by RNA polymerase II is limited by melting of the promoter DNA in the region immediately upstream of the initiation site. In *The Journal of biological chemistry* 269 (48), pp. 30101–30104.

Pardee, T. S.; Bangur, C. S.; Ponticelli, A. S. (1998): The N-terminal region of yeast TFIIB contains two adjacent functional domains involved in stable RNA polymerase II binding and transcription start site selection. In *The Journal of biological chemistry* 273 (28), pp. 17859–17864.

Parvin, J. D.; Sharp, P. A. (1993): DNA topology and a minimal set of basal factors for transcription by RNA polymerase II. In *Cell* 73 (3), pp. 533–540.

Pasman, Z.; Hippel, P. H. von (2000): Regulation of rho-dependent transcription termination by NusG is specific to the Escherichia coli elongation complex. In *Biochemistry* 39 (18), pp. 5573–5585.

Paytubi, Sonia; White, Malcolm F. (2009): The crenarchaeal DNA damage-inducible transcription factor B paralogue TFB3 is a general activator of transcription. In *Molecular microbiology* 72 (6), pp. 1487–1499. DOI: 10.1111/j.1365-2958.2009.06737.x.

Peeters, Eveline; Charlier, Daniel (2010): The Lrp family of transcription regulators in archaea. In *Archaea (Vancouver, B.C.)* 2010, p. 750457. DOI: 10.1155/2010/750457.

Peeters, Eveline; Driessen, Rosalie P. C.; Werner, Finn; Dame, Remus T. (2015): The interplay between nucleoid organization and transcription in archaeal genomes. In *Nature reviews. Microbiology* 13 (6), pp. 333–341. DOI: 10.1038/nrmicro3467.

Peng, Nan; Xia, Qiu; Chen, Zhengjun; Liang, Yun Xiang; She, Qunxin (2009): An upstream activation element exerting differential transcriptional activation on an archaeal promoter. In *Molecular microbiology* 74 (4), pp. 928–939. DOI: 10.1111/j.1365-2958.2009.06908.x.

Pinto, I.; Ware, D. E.; Hampsey, M. (1992): The yeast SUA7 gene encodes a homolog of human transcription factor TFIIB and is required for normal start site selection in vivo. In *Cell* 68 (5), pp. 977–988.

Plaschka, C.; Hantsche, M.; Dienemann, C.; Burzinski, C.; Plitzko, J.; Cramer, P. (2016): Transcription initiation complex structures elucidate DNA opening. In *Nature* 533 (7603), pp. 353–358. DOI: 10.1038/nature17990.

Pluchon, Pierre-Francois; Fouqueau, Thomas; Creze, Christophe; Laurent, Sebastien; Briffotaux, Julien; Hogrel, Gaelle et al. (2013): An extended network of genomic maintenance in the archaeon Pyrococcus abyssi highlights unexpected associations between eucaryotic homologs. In *PloS one* 8 (11), e79707. DOI: 10.1371/journal.pone.0079707.

Ponjavic, Jasmina; Lenhard, Boris; Kai, Chikatoshi; Kawai, Jun; Carninci, Piero; Hayashizaki, Yoshihide; Sandelin, Albin (2006): Transcriptional and structural impact of TATA-initiation site spacing in mammalian core promoters. In *Genome biology* 7 (8), R78. DOI: 10.1186/gb-2006-7-8-R78.

Porrua, Odil; Boudvillain, Marc; Libri, Domenico (2016): Transcription Termination: Variations on Common Themes. In *Trends in genetics : TIG* 32 (8), pp. 508–522. DOI: 10.1016/j.tig.2016.05.007.

Preus, Soren; Noer, Sofie L.; Hildebrandt, Lasse L.; Gudnason, Daniel; Birkedal, Victoria (2015): iSMS: single-molecule FRET microscopy software. In *Nature methods* 12 (7), pp. 593–594. DOI: 10.1038/nmeth.3435.

Qureshi, S. A.; Bell, S. D.; Jackson, S. P. (1997): Factor requirements for transcription in the Archaeon Sulfolobus shibatae. In *The EMBO journal* 16 (10), pp. 2927–2936. DOI: 10.1093/emboj/16.10.2927.

Reeve, J. N.; Bailey, K. A.; Li, W-T; Marc, F.; Sandman, K.; Soares, D. J. (2004): Archaeal histones: structures, stability and DNA binding. In *Biochemical Society transactions* 32 (Pt 2), pp. 227–230.

Reeve, John N. (2003): Archaeal chromatin and transcription. In *Molecular microbiology* 48 (3), pp. 587–598.

Reich, Christoph; Zeller, Mirijam; Milkereit, Philipp; Hausner, Winfried; Cramer, Patrick; Tschochner, Herbert; Thomm, Michael (2009): The archaeal RNA polymerase subunit P and the eukaryotic polymerase subunit Rpb12 are interchangeable in vivo and in vitro. In *Molecular microbiology* 71 (4), pp. 989–1002. DOI: 10.1111/j.1365-2958.2008.06577.x.

Reichelt, Robert; Gindner, Antonia; Thomm, Michael; Hausner, Winfried (2016): Genome-wide binding analysis of the transcriptional regulator TrmBL1 in Pyrococcus furiosus. In *BMC genomics* 17, p. 40. DOI: 10.1186/s12864-015-2360-0.

Reikofski, J.; Tao, B. Y. (1992): Polymerase chain reaction (PCR) techniques for site-directed mutagenesis. In *Biotechnology advances* 10 (4), pp. 535–547.

Renfrow, Matthew B.; Naryshkin, Nikolai; Lewis, L. Michelle; Chen, Hung-Ta; Ebright, Richard H.; Scott, Robert A. (2004): Transcription factor B contacts promoter DNA near the transcription start site of the archaeal transcription initiation complex. In *The Journal of biological chemistry* 279 (4), pp. 2825–2831. DOI: 10.1074/jbc.M311433200.

Revyakin, Andrey; Liu, Chenyu; Ebright, Richard H.; Strick, Terence R. (2006): Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. In *Science (New York, N.Y.)* 314 (5802), pp. 1139–1143. DOI: 10.1126/science.1131398.

Richard, Derek J.; Bell, Stephen D.; White, Malcolm F. (2004): Physical and functional interaction of the archaeal single-stranded DNA-binding protein SSB with RNA polymerase. In *Nucleic acids research* 32 (3), pp. 1065–1074. DOI: 10.1093/nar/gkh259.

Ruprich-Robert, Gwenael; Thuriaux, Pierre (2010): Non-canonical DNA transcription enzymes and the conservation of two-barrel RNA polymerases. In *Nucleic acids research* 38 (14), pp. 4559–4569. DOI: 10.1093/nar/gkq201.

Sainsbury, Sarah; Niesser, Jurgen; Cramer, Patrick (2013): Structure and function of the initially transcribing RNA polymerase II-TFIIB complex. In *Nature* 493 (7432), pp. 437–440. DOI: 10.1038/nature11715.

Santangelo, Thomas J.; Cubonova, L'ubomira; Skinner, Katherine M.; Reeve, John N. (2009): Archaeal intrinsic transcription termination in vivo. In *Journal of bacteriology* 191 (22), pp. 7102–7108. DOI: 10.1128/JB.00982-09.

Santangelo, Thomas J.; Reeve, John N. (2006): Archaeal RNA polymerase is sensitive to intrinsic termination directed by transcribed and remote sequences. In *Journal of molecular biology* 355 (2), pp. 196–210. DOI: 10.1016/j.jmb.2005.10.062.

Sassone-Corsi, P.; Corden, J.; Kedinger, C.; Chambon, P. (1981): Promotion of specific in vitro transcription by excised "TATA" box sequences inserted in a foreign nucleotide environment. In *Nucleic acids research* 9 (16), pp. 3941–3958.

Sauer, F.; Fondell, J. D.; Ohkuma, Y.; Roeder, R. G.; Jackle, H. (1995): Control of transcription by Kruppel through interactions with TFIIB and TFIIE beta. In *Nature* 375 (6527), pp. 162–164. DOI: 10.1038/375162a0.

Schaeffer, L.; Roy, R.; Humbert, S.; Moncollin, V.; Vermeulen, W.; Hoeijmakers, J. H. et al. (1993): DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. In *Science (New York, N.Y.)* 260 (5104), pp. 58–63.

Schulz, Sarah; Gietl, Andreas; Smollett, Katherine; Tinnefeld, Philip; Werner, Finn; Grohmann, Dina (2016): TFE and Spt4/5 open and close the RNA polymerase clamp during the transcription cycle. In *Proceedings of the National Academy of Sciences of the United States of America* 113 (13), E1816-25. DOI: 10.1073/pnas.1515817113.

Shapiro, James A. (2009): Revisiting the central dogma in the 21st century. In *Annals of the New York Academy of Sciences* 1178, pp. 6–28. DOI: 10.1111/j.1749-6632.2009.04990.x.

Shaw, S. P.; Wingfield, J.; Dorsey, M. J.; Ma, J. (1996): Identifying a species-specific region of yeast TF11B in vivo. In *Molecular and cellular biology* 16 (7), pp. 3651–3657.

Shockley, Keith R.; Ward, Donald E.; Chhabra, Swapnil R.; Conners, Shannon B.; Montero, Clemente I.; Kelly, Robert M. (2003): Heat shock response by the hyperthermophilic archaeon Pyrococcus furiosus. In *Applied and environmental microbiology* 69 (4), pp. 2365–2371.

Sievers, Fabian; Wilm, Andreas; Dineen, David; Gibson, Toby J.; Karplus, Kevin; Li, Weizhong et al. (2011): Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. In *Molecular systems biology* 7, p. 539. DOI: 10.1038/msb.2011.75.

Sikorski, Timothy W.; Ficarro, Scott B.; Holik, John; Kim, TaeSoo; Rando, Oliver J.; Marto, Jarrod A.; Buratowski, Stephen (2011): Sub1 and RPA associate with RNA polymerase II at different stages of transcription. In *Molecular cell* 44 (3), pp. 397–409. DOI: 10.1016/j.molcel.2011.09.013.

Skowyra, Agnieszka; MacNeill, Stuart A. (2012): Identification of essential and non-essential single-stranded DNA-binding proteins in a model archaeal organism. In *Nucleic acids research* 40 (3), pp. 1077–1090. DOI: 10.1093/nar/gkr838.

Soares, D.; Dahlke, I.; Li, W. T.; Sandman, K.; Hethke, C.; Thomm, M.; Reeve, J. N. (1998): Archaeal histone stability, DNA binding, and transcription inhibition above 90 degrees C. In *Extremophiles : life under extreme conditions* 2 (2), pp. 75–81.

Sommer, Bettina; Waege, Ingrid; Pollmann, David; Seitz, Tobias; Thomm, Michael; Sterner, Reinhard; Hausner, Winfried (2014): Activation of a chimeric Rpb5/RpoH subunit using library selection. In *PloS one* 9 (1), e87485. DOI: 10.1371/journal.pone.0087485.

Soppa, J. (1999): Transcription initiation in Archaea: facts, factors and future aspects. In *Molecular microbiology* 31 (5), pp. 1295–1305.

Sosunov, Vasily; Sosunova, Ekaterina; Mustaev, Arkady; Bass, Irina; Nikiforov, Vadim; Goldfarb, Alex (2003): Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. In *The EMBO journal* 22 (9), pp. 2234–2244. DOI: 10.1093/emboj/cdg193.

Spitalny, Patrizia; Thomm, Michael (2003): Analysis of the open region and of DNA-protein contacts of archaeal RNA polymerase transcription complexes during transition from initiation to elongation. In *The Journal of biological chemistry* 278 (33), pp. 30497–30505. DOI: 10.1074/jbc.M303633200.

Spitalny, Patrizia; Thomm, Michael (2008): A polymerase III-like reinitiation mechanism is operating in regulation of histone expression in archaea. In *Molecular microbiology* 67 (5), pp. 958–970. DOI: 10.1111/j.1365-2958.2007.06084.x.

T.A. Hall: BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. In *Nucleic Acids Symposium Series* 1999 (41), pp. 95–98. Available online at http://www.mbio.ncsu.edu/bioedit/page2.html.

Tan, S.; Hunziker, Y.; Sargent, D. F.; Richmond, T. J. (1996): Crystal structure of a yeast TFIIA/TBP/DNA complex. In *Nature* 381 (6578), pp. 127–151. DOI: 10.1038/381127a0.

Tanaka, Aki; Akimoto, Yusuke; Kobayashi, Satoko; Hisatake, Koji; Hanaoka, Fumio; Ohkuma, Yoshiaki (2015): Association of the winged helix motif of the TFIIEalpha subunit of TFIIE with either the TFIIEbeta subunit or TFIIB distinguishes its functions in transcription. In *Genes to cells : devoted to molecular & cellular mechanisms* 20 (3), pp. 203–216. DOI: 10.1111/gtc.12212.

Thomm, M. (1996): Archaeal transcription factors and their role in transcription initiation. In *FEMS microbiology reviews* 18 (2-3), pp. 159–171.

Thomm, M.; Wich, G. (1988): An archaebacterial promoter element for stable RNA genes with homology to the TATA box of higher eukaryotes. In *Nucleic acids research* 16 (1), pp. 151–163.

Thompson, D. K.; Palmer, J. R.; Daniels, C. J. (1999): Expression and heat-responsive regulation of a TFIIB homologue from the archaeon Haloferax volcanii. In *Molecular microbiology* 33 (5), pp. 1081–1092.

Tirode, F.; Busso, D.; Coin, F.; Egly, J. M. (1999): Reconstitution of the transcription factor TFIIH: assignment of functions for the three enzymatic subunits, XPB, XPD, and cdk7. In *Molecular cell* 3 (1), pp. 87–95.

Tokusumi, Yumiko; Ma, Ying; Song, Xianzhou; Jacobson, Raymond H.; Takada, Shinako (2007): The new core promoter element XCPE1 (X Core Promoter Element 1) directs activator-, mediator-, and TATA-binding protein-dependent but TFIID-independent RNA polymerase II transcription from TATA-less promoters. In *Molecular and cellular biology* 27 (5), pp. 1844–1858. DOI: 10.1128/MCB.01363-06.

Tran, Khiem; Gralla, Jay D. (2008): Control of the timing of promoter escape and RNA catalysis by the transcription factor IIb fingertip. In *The Journal of biological chemistry* 283 (23), pp. 15665–15671. DOI: 10.1074/jbc.M801439200.

Tschochner, H.; Sayre, M. H.; Flanagan, P. M.; Feaver, W. J.; Kornberg, R. D. (1992): Yeast RNA polymerase II initiation factor e: isolation and identification as the functional counterpart of human transcription factor IIB. In *Proceedings of the National Academy of Sciences of the United States of America* 89 (23), pp. 11292–11296.

Vannini, Alessandro; Cramer, Patrick (2012): Conservation between the RNA polymerase I, II, and III transcription initiation machineries. In *Molecular cell* 45 (4), pp. 439–446. DOI: 10.1016/j.molcel.2012.01.023.

Vassylyev, Dmitry G.; Vassylyeva, Marina N.; Perederina, Anna; Tahirov, Tahir H.; Artsimovitch, Irina (2007): Structural basis for transcription elongation by bacterial RNA polymerase. In *Nature* 448 (7150), pp. 157–162. DOI: 10.1038/nature05932.

Wang, Z.; Roeder, R. G. (1997): Three human RNA polymerase III-specific subunits form a subcomplex with a selective function in specific transcription initiation. In *Genes & development* 11 (10), pp. 1315–1326.

Wardleworth, B. N.; Russell, R. J. M.; Bell, S. D.; Taylor, G. L.; White, M. F. (2002): Structure of Alba: an archaeal chromatin protein modulated by acetylation. In *The EMBO journal* 21 (17), pp. 4654–4662.

Washburn, Robert S.; Gottesman, Max E. (2015): Regulation of transcription elongation and termination. In *Biomolecules* 5 (2), pp. 1063–1078. DOI: 10.3390/biom5021063.

WATSON, J. D.; Crick, F. (1953): Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. In *Nature* 171 (4356), pp. 737–738.

Weinzierl, Robert O. J. (2011): The Bridge Helix of RNA polymerase acts as a central nanomechanical switchboard for coordinating catalysis and substrate movement. In *Archaea (Vancouver, B.C.)* 2011, p. 608385. DOI: 10.1155/2011/608385.

Wenzel, Sabine; Martins, Berta M.; Rosch, Paul; Wohrl, Birgitta M. (2010): Crystal structure of the human transcription elongation factor DSIF hSpt4 subunit in complex with the hSpt5 dimerization interface. In *The Biochemical journal* 425 (2), pp. 373–380. DOI: 10.1042/BJ20091422.

Werner, Finn (2007): Structure and function of archaeal RNA polymerases. In *Molecular microbiology* 65 (6), pp. 1395–1404. DOI: 10.1111/j.1365-2958.2007.05876.x.

Werner, Finn (2008): Structural evolution of multisubunit RNA polymerases. In *Trends in microbiology* 16 (6), pp. 247–250. DOI: 10.1016/j.tim.2008.03.008.

Werner, Finn (2012): A nexus for gene expression-molecular mechanisms of Spt5 and NusG in the three domains of life. In *Journal of molecular biology* 417 (1-2), pp. 13–27. DOI: 10.1016/j.jmb.2012.01.031.

Werner, Finn; Grohmann, Dina (2011): Evolution of multisubunit RNA polymerases in the three domains of life. In *Nature reviews. Microbiology* 9 (2), pp. 85–98. DOI: 10.1038/nrmicro2507.

Werner, Finn; Weinzierl, Robert O. J. (2005): Direct modulation of RNA polymerase core functions by basal transcription factors. In *Molecular and cellular biology* 25 (18), pp. 8344–8355. DOI: 10.1128/MCB.25.18.8344-8355.2005.

Wettach, J.; Gohl, H. P.; Tschochner, H.; Thomm, M. (1995): Functional interaction of yeast and human TATA-binding proteins with an archaeal RNA polymerase and promoter. In *Proceedings of the National Academy of Sciences of the United States of America* 92 (2), pp. 472–476.

Wiesler, Simone C.; Weinzierl, Robert O. J. (2011): The linker domain of basal transcription factor TFIIB controls distinct recruitment and transcription stimulation functions. In *Nucleic acids research* 39 (2), pp. 464–474. DOI: 10.1093/nar/gkq809.

Wild, Thomas; Cramer, Patrick (2012): Biogenesis of multisubunit RNA polymerases. In *Trends in biochemical sciences* 37 (3), pp. 99–105. DOI: 10.1016/j.tibs.2011.12.001.

Wilkinson, Steven P.; Ouhammouch, Mohamed; Geiduschek, E. Peter (2010): Transcriptional activation in the context of repression mediated by archaeal histones. In *Proceedings of the National Academy of Sciences of the United States of America* 107 (15), pp. 6777–6781. DOI: 10.1073/pnas.1002360107.

Willis, Ian M. (2002): A universal nomenclature for subunits of the RNA polymerase III transcription initiation factor TFIIIB. In *Genes & development* 16 (11), pp. 1337–1338.

Witkowski, Leora; Foulkes, William D. (2015): In Brief: Picturing the complex world of chromatin remodelling families. In *The Journal of pathology* 237 (4), pp. 403–406. DOI: 10.1002/path.4585.

Woese, C. R.; Kandler, O.; Wheelis, M. L. (1990): Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. In *Proceedings of the National Academy of Sciences of the United States of America* 87 (12), pp. 4576–4579.

Wojtas, Magdalena; Peralta, Bibiana; Ondiviela, Marina; Mogni, Maria; Bell, Stephen D.; Abrescia, Nicola G. A. (2011): Archaeal RNA polymerase: the influence of the protruding stalk in crystal packing and preliminary biophysical analysis of the Rpo13 subunit. In *Biochemical Society transactions* 39 (1), pp. 25–30. DOI: 10.1042/BST0390025.

Xia, Yongzhen; Chu, Wenqiao; Qi, Qingsheng; Xun, Luying (2015): New insights into the QuikChange process guide the use of Phusion DNA polymerase for site-directed mutagenesis. In *Nucleic acids research* 43 (2), e12. DOI: 10.1093/nar/gku1189.

Xie, Yunwei; Reeve, John N. (2004a): Transcription by an archaeal RNA polymerase is slowed but not blocked by an archaeal nucleosome. In *Journal of bacteriology* 186 (11), pp. 3492–3498. DOI: 10.1128/JB.186.11.3492-3498.2004.

Xie, Yunwei; Reeve, John N. (2004b): Transcription by Methanothermobacter thermautotrophicus RNA polymerase in vitro releases archaeal transcription factor B but not TATA-box binding protein from the template DNA. In *Journal of bacteriology* 186 (18), pp. 6306–6310. DOI: 10.1128/JB.186.18.6306-6310.2004.

Xing, L.; Gopal, V. K.; Quinn, P. G. (1995): cAMP response element-binding protein (CREB) interacts with transcription factors IIB and IID. In *The Journal of biological chemistry* 270 (29), pp. 17488–17493.

Yan, Q.; Moreland, R. J.; Conaway, J. W.; Conaway, R. C. (1999): Dual roles for transcription factor IIF in promoter escape by RNA polymerase II. In *The Journal of biological chemistry* 274 (50), pp. 35668–35675.

Young, Travis S.; Ahmad, Insha; Yin, Jun A.; Schultz, Peter G. (2010): An enhanced system for unnatural amino acid mutagenesis in E. coli. In *Journal of molecular biology* 395 (2), pp. 361–374. DOI: 10.1016/j.jmb.2009.10.030.

Zeller, Mirijam E. (2011): Die Funktion von Transkriptionsfaktor (II) B und Mechanismus der Transkriptionsinitiation. Available online at http://epub.uni-regensburg.de/18671/.

Zhang, G.; Campbell, E. A.; Minakhin, L.; Richter, C.; Severinov, K.; Darst, S. A. (1999): Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 A resolution. In *Cell* 98 (6), pp. 811–824.

Zhang, Yu; Feng, Yu; Chatterjee, Sujoy; Tuske, Steve; Ho, Mary X.; Arnold, Eddy; Ebright, Richard H. (2012a): Structural basis of transcription initiation. In *Science (New York, N.Y.)* 338 (6110), pp. 1076–1080. DOI: 10.1126/science.1227786.

Zhang, ZhenFeng; Guo, Li; Huang, Li (2012b): Archaeal chromatin proteins. In *Science China. Life sciences* 55 (5), pp. 377–385. DOI: 10.1007/s11427-012-4322-y.

Zhang, Zhengjian; English, Brian P.; Grimm, Jonathan B.; Kazane, Stephanie A.; Hu, Wenxin; Tsai, Albert et al. (2016): Rapid dynamics of general transcription factor TFIIB binding during preinitiation complex assembly revealed by single-molecule analysis. In *Genes & development* 30 (18), pp. 2106–2118. DOI: 10.1101/gad.285395.116.

Zhu, W.; Zeng, Q.; Colangelo, C. M.; Lewis, M.; Summers, M. F.; Scott, R. A. (1996): The N-terminal domain of TFIIB from Pyrococcus furiosus forms a zinc ribbon. In *Nature structural biology* 3 (2), pp. 122–124.

Ziegler, Lynn M.; Khaperskyy, Denys A.; Ammerman, Michelle L.; Ponticelli, Alfred S. (2003): Yeast RNA polymerase II lacking the Rpb9 subunit is impaired for interaction with transcription factor IIF. In *The Journal of biological chemistry* 278 (49), pp. 48950–48956. DOI: 10.1074/jbc.M309656200.

X.   Danksagung

Ich bedanke mich recht herzlich bei Herrn Prof. Michael Thomm für die Vergabe dieses interessanten Themas, und der Möglichkeit, an seinem Lehrstuhl zu promovieren. Das von Ihm entgegengebrachte Vertrauen und seine Unterstützung habe ich immer sehr geschätzt.

Mein Weiterer Dank geht an Dr. Winfried Hausner, und an Prof. Dina Grohmann, die mich in allen Angelegenheiten immer unterstützt haben, und mir mit Rat zur Seite standen.

Die Atmosphäre am Lehrstuhl war immer sehr gut, und dafür möchte ich allen Studenten, den Assistenten Wolfgang, Renate, und Gabi, den Kellerkindern Konni, Thomas und Leonardo herzlich danken.

Bei meinen Kollegen Felix Grünberger, sowie ehemalige Doktoranden des Lehrstuhls, Ingrid Waege, Antonia Gindner, Lydia Kreuter, Stefanie Daxer, Pia Wiegmann und Franklin Vincent möchte ich mich für die angeregten Diskussionen und die gemeinsame Zeit am Lehrstuhl bedanken. Insbesondere danke ich meinem Kollegen Dr. Robert Reichelt, für so ziemlich alles.

Ein ganz besonderer Dank geht an das „Herz" des Lehrstuhls, Elisabeth Nagelfeld. Ohne ihr Engagement und Unterstützung wäre vieles nicht möglich gewesen.

Ich danke meinen Freunden, Bekannten und meiner ganzen Familie für die Unterstützung und für die geistige Wiederbelebung in schweren Zeiten. Insbesondere möchte ich mich bei meiner Frau Andrea, die mich in all den Jahren immer voll unterstützt und die aufreibende Zeit des Promotionsstudiums erduldet hat, ganz herzlich bedanken. Gerade in der Schlussphase warst du immer mein Ankerpunkt und für mich da. Meine Tochter Victoria, der Mittelpunkt meines Lebens, hat mir ebenfalls immer die Kraft und Motivation gegeben, meine Arbeit zu Ende zu bringen.

„In hundert Jahren in alles vorbei".

XI.   Erklärung

Ich versichere hiermit, dass ich die Arbeit selbständig angefertigt habe und keine anderen als die hier angegebenen Quellen und Hilfsmittel benutzt sowie die wörtlich oder inhaltlich übernommenen Stellen als solche kenntlich gemacht habe. Diese Arbeit war bisher noch nicht Bestandteil eines Prüfungsverfahrens, andere Promotionsgesuche wurden nicht unternommen.

Neumarkt, den 18.11.2016

....

Dexl Stefan