Summer 9-18-2015

# The Role of *Pyrococcus furiosus* Transcription Factor E in Transcription Iniitiation

Robyn Lynn Eustis
*Portland State University*

# The Role of *Pyrococcus furiosus* Transcription Factor E in Transcription Initiation

by

Robyn Lynn Eustis

A thesis submitted in partial fulfillment of the
requirements for the degree of

Master of Science
in
Biology

Thesis Committee:
Michael Bartlett, Chair
Jeffrey Singer
Kenneth Stedman

Portland State University
2015

ABSTRACT

All sequenced archaeal genomes encode a general transcription factor, TFE, which is highly conserved and homologous to the alpha subunit of the eukaryotic transcription factor TFIIE. TFE functions to increase promoter opening efficiency during transcription initiation, although the mechanism for this is unclear. The N-terminus of TFE contains a common DNA binding motif, a winged helix. At the tip of this winged helix is a highly conserved region of aromatic amino acids that is close to DNA during initiation. TFE activation can compensate for mutations in another transcription factor, TFB2, which is homologous to TFIIB. *P. furiosus* encodes two paralogs of the eukaryotic RNA polymerase II transcription factor TFIIB: TFB1 and TFB2. TFB2 lacks a portion of the highly conserved N-terminus, and functions in transcription complexes at a lower efficiency than TFB1. It has been demonstrated that the presence of TFE is able to assist in transcription with TFB2 *in vitro* bringing its efficiency to almost TFB1 levels. Thus, TFB2 provides a unique opportunity to evaluate the function of the TFE winged helix in transcription. In this study the aromatic patch of the TFE winged helix was mutated to test its role in activation of TFB1 and TFB2-containing transcription complexes, because this aromatic patch is required for full TFE activity especially when NTP concentrations are low.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my husband Jeff Eustis and my children Ashley and Sean. Without their support I would not have been able to accomplish this goal. Thank you to my parents. I would also like to express my gratitude to my advisor Dr. Michael Bartlett for his guidance and encouragement throughout this process. Thank you also to my committee members Dr. Kenneth Stedman and Dr. Jeffrey Singer

**TABLE OF CONTENTS**

## LIST OF TABLES

**LIST OF FIGURES**

# CHAPTER 1

## 1. INTRODUCTION

The universal tree of life is divided into three branches: Eukarya, Bacteria, and Archaea. Archaea and bacteria are prokaryotes; they both lack a nuclear envelope and membrane-bound organelles, while eukaryotes contain their genetic material in a nucleus and have separate organelles performing specific functions within the cell. As the number of sequenced archaeal genomes has increased, it has become clear that archaea contain a combination of bacterial and eukaryotic features. Metabolic and structural genes are most similar to those of bacteria. However, archaeal genes involving information processing, such as replication, transcription, and translation show striking similarity to their eukaryotic counterparts *(Soppa 1999),* although in archaea these processes tend to be "scaled down", with fewer factors involved. Based on these fundamental similarities, the archaeal system provides a simple model for understanding the eukaryotic transcription mechanism and its evolution.

Many of the known archaea are hyperthermophiles, organisms that flourish at high temperatures, growing optimally at or above 80°C (*Stetter, 2006).* It has been proposed that the last common ancestor was a hyperthermophile based on its position in the tree of life *(Stetter, 2006)*. Therefore, study of archaeal hyperthermophiles such as *Pyrococcus furiosus* the organism used here, could provide a window into early evolution and the determinants of transcription at high temperature.

**1.1 Transcription in the three domains of life.**

Transcription is the process whereby RNA is synthesized from a DNA template. It is the first step in gene expression and is catalyzed by the enzyme, RNA polymerase, in conjunction with other transcription factors. Transcription in archaea, while scaled down, is fundamentally homologous to eukarya with central components of the process being highly conserved between the two groups *(Hausner, Wettach, Hethke, & Thomm, 1996)*. Specifically, archaeal transcription is most similar to the eukaryotic RNA polymerase II (RNAP II) transcriptional apparatus *(Thomm, 1996)*.

Eukaryotes have diversified their synthesis of RNAs and have three separate nuclear RNA polymerases, I, II, and III, each transcribing a particular class of RNA. Archaea like bacteria maintain only one RNA polymerase. Archaeal RNA polymerase is structurally similar to that of eukaryotic RNA polymerase II as demonstrated by comparison of the structure of the RNA polymerase from the archaeon *Sulfolobus shibatae*. with the structure of *Saccharomyces cerevisiae* RNA polymerase (*Hirata, Klein, & Murakami, 2008*). There are 13 subunits demonstrated in the archaeal RNA polymerase and many of these share structural, functional and sequence similarity with RNAP II subunits *(Korkhin et al., 2009)*. In contrast, bacterial transcription varies from that of eukaryal and archaeal in that the polymerase is smaller and has fewer subunits. It is composed of the core enzyme, which contains five subunits, and the sigma subunit which when joined to the core enzyme forms the holoenzyme *(Molodtsov et al., 2013)*. In bacterial transcription a sigma factor is required for transcription start site recognition, and typically no other accessory factors are required for transcription initiation in the bacterial system. In

eukaryotes, RNA polymerase II transcription requires a host of other factors, including TFIID, which contains TBP (TATA box binding protein), TFIIB, TFIIE, TFIIH, TFIIF, and TFIIA (*Grunberg & Hahn, 2013*) (see figure 1).

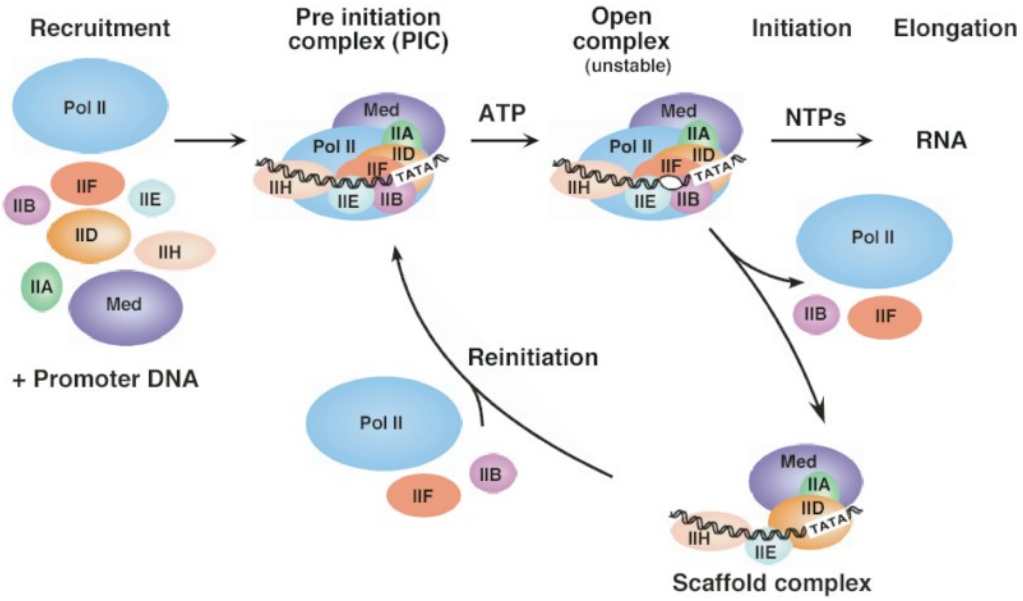



Figure 1: Schematic illustrating the assembly of transcription factors in eukaryotic transcription. *(Hahn, 2005)*

Archaeal genomes contain several transcription factors that are orthologous to eukaryotic general transcription factors and display high levels of sequence conservation. These factors are TBP, TFB, and TFE. Archaeal TBP is homologous to TBP in eukaryotes. It is essential for transcription and is responsible for recognition of the TATA box *(DeDecker et al., 1996)*. TFB is homologous to eukaryotic TFIIB and is also essential for transcription and is responsible for promoter recognition and transcription start site selection *(Littlefield, Korkhin, & Sigler, 1999)*. TFE is homologous to the N- terminal portion of the alpha subunit of TFIIE *(Meinhart, Blobel,*

*& Cramer, 2003)*. TFE has been shown to stimulate transcription *in vitro* under low TBP conditions or in the presence of a secondary, low activity TFB *(Micorescu et al., 2008)*. Some archaeal genomes such as *Halobacterium* encode multiple TBPs and multiple TFBs, implying that differing combinations of the two could be responsible for transcription programming under varied conditions *(Bleiholder, Frommherz, Teufel, & Pfeifer, 2012)*. *Thermococcus kodakarensis* encodes two TFB proteins, either of which is sufficient for basal transcription *(Santangelo, Cubonova, James, & Reeve, 2007)*. *Pyrococcus furiosus* also encodes two TFB proteins, designated TFB1 and TFB2. TFB1 is highly active in transcription and is highly homologous to TFIIB throughout its length. The N-terminus if TFB1 contains a region known as the B-reader which is responsible for transcription start site selection *(Kostrewa et al., 2009)*. The C-terminus of TFB2 is 63% identical to TFB1 and is a well-conserved helix-turn-helix motif whose role is recognition of the B recognition element (BRE) located on the DNA upstream of the transcription start site.

The N-terminus of TFB2, however, is not well conserved and contains only 45% identity to the N-terminus of TFB1. TFB2 lacks residues in the N-terminal B-reader conserved sequences and *in vitro* transcription assays demonstrate decreased transcription *(Micorescu et al., 2008)*(See figure 2). There are no known homologs to TFIIA, TFIIH, or TFIIF found in archaea *(Soppa & Universitèat Frankfurt, 1999)*. *In vitro* transcription with purified archaeal RNAP requires only TBP and TFB. This minimal system provides an unencumbered view of transcription as compared to eukarya and can provide insight into the more complex eukaryotic-type transcriptional process.

Figure 2A.

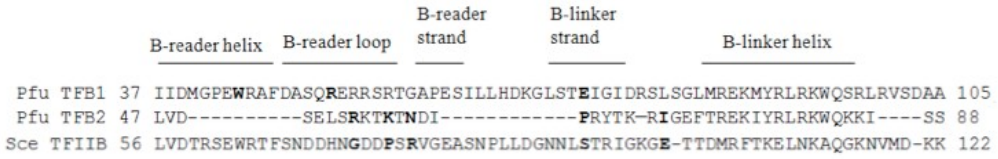



Figure 2B.

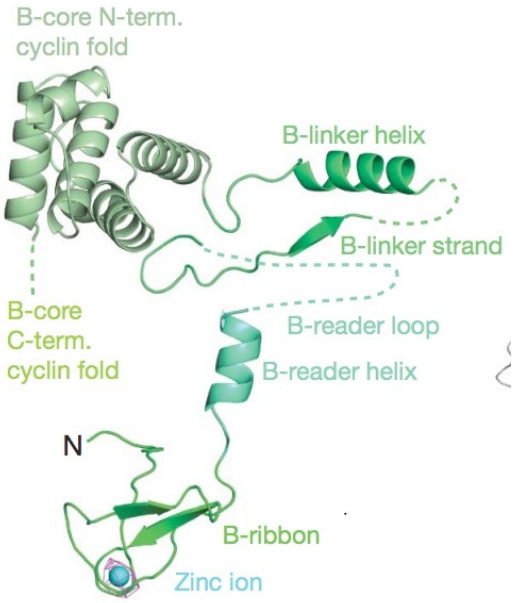



Figure 2: A. Manual sequence alignment of TFBs from *P. furiosus* and yeast *(Bhattarai, Thesis 2013)*. B. Model of yeast TFIIB demonstrating key regions in transcription *(Kostrewa et al., 2009)*.

## 1.2 Promoter opening in transcription initiation

Transcription in all domains of life is accomplished by RNAP through a series of steps: Initiation, elongation and termination. During initiation, transcription factors assist RNAP in binding to a promoter and the promoter is opened and the first few nucleotides of the transcript are produced. Elongation begins when the transcript extends past ~ 13 nucleotides and RNAP leaves the promoter. Upon termination the RNA transcript is released and RNAP dissociates from DNA *(Debnath, Roy, Bera, Ghoshal, & Roy, 2013).*

### 1.2.1 Promoter opening by RNAP II

Transcription in eukaryotes requires the presence of multiple general transcription factors in addition to the RNA polymerase that contribute various functions during the initiation pathway. The preinitiation complex is formed by the interactions of TFIIA, TFIIB, TBP (as part of TFIID), TFIIF, TFIIH, and TFIIE with RNAP II on the duplex DNA (closed complex) *(He, Fang, Taatjes, & Nogales, 2013).* In eukaryotes transcription factors interact at the promoter region in a stepwise manner. The 14 subunit TFIID is the first general transcription factor to bind the promoter, through interaction of TBP with the TATA box sequence of the core promoter. TBP binds at the TATA box in the minor groove and creates a substantial bend in the DNA *(Gietl et al., 2014)*. This provides a platform for the remaining transcription factors to bind. TFIIA

joins the complex and assists in stabilizing the nascent PIC by interacting with TBP and DNA upstream of the TATA region *(He et al., 2013)*. TFIIB binds to TBP and to DNA upstream and downstream of TBP further stabilizing the initiation complex *(He et al., 2013)*. RNAP is recruited through specific interactions between functional domains of TFIIB and the polymerase *(He et al., 2013)*. The TFIIB B-ribbon binds to the Pol II dock domain. The B-reader helix enters the Pol II active site cleft where it positions the B- reader loop on the Pol II rudder. The B-linker binds to the Pol II clamp coiled-coil where it interacts with the TFIIB N-terminal cyclin repeat at the Pol II wall *(Grunberg, Warfield, & Hahn, 2012)*. TFIIF then joins the promoter in associatiation with the polymerase *(He et al., 2013)*. The effect of TFIIF on the pre-initiation complex (PIC) is to stabilize DNA in the Pol II cleft *(Grunberg et al., 2012)*. Recent cryo-EM analysis reveals that addition of TFIIF results in clamp opening and positioning of the promoter over the Pol II cleft *(He et al., 2013)*. TFIIE consists of two subunits: the N-terminal alpha subunit and the C-terminal beta subunit *(Grunberg and Hahn, 2013*). The PIC is stabilized through interactions between TFIIE C-terminus and the pol II stalk region as well as the TFIIE N-terminus and the RNAP clamp coiled coil *(Grunberg and Hahn, 2013*). TFIIE then recruits TFIIH to the assemblage and together they assist RNAP in forming the open complex in preparation for transcription *(Grunberg and Hahn, 2013*). TFIIH contains 10 subunits of which one is an ATP dependent helicase, XPB, which is responsible for creating a 10-12 basepair transcription bubble (open complex). The transcription start site is then translocated to the active site of pol II and initiation commences (*Grunberg and Hahn 2013)*.

### 1.2.2. Promoter opening by archael RNAP

The archaeal transcription machinery is highly similar to eukaryotic transcription with respect to the structure and function of the RNA polymerase, but requires fewer factors *(Werner & Grohmann, 2011)*. The archaeal RNA polymerase contains 13 subunits that are homologous in both structure and function to those of RNA polymerase II in eukaryotes. *(Werner & Grohmann, 2011)* Three general transcription factors assist in transcription initiation: TBP, TFB, and TFE which are homologous to eukaryotic TBP, TFIIB, and the TFIIE alpha subunit respectively *(Werner & Grohmann, 2011)*. In a similar manner to that observed in eukaryotes TBP binds the TATA box in the archaeal promoter region distorting the DNA to an angle of roughly 75-80 degrees *(Littlefield et al., 1999)*. TFB then binds to the B-recognition element (BRE) located upstream of the TATA box. The DNA bound TBP/TFB subcomplex recruits RNA polymerase forming the preinitiation complex (PIC). Transcription factor E (TFE) associates with RNAP and assists in promoter melting, as well as stabilizing the PIC through interaction with the non-transcribed strand (NTS) *(Werner & Grohmann, 2011)*.

### 1.2.3 Promoter opening by bacterial RNAP

Components of bacterial transcription are: the core RNA polymerase, consisting of two identical alpha ($\alpha$) subunits, one beta ($\beta$), one $\beta$', an omega ($\omega$) subunit, and the sigma ($\sigma$) subunit, which is responsible for promoter recognition. (*Feklistov and Darst, 2011)*. The RNA polymerase holoenzyme (the RNA polymerase and the associated

sigma factor) resembles a crab claw with β and β'subunits forming the pincers. It is within this structure that the catalytic $Mg^{2+}$ resides, forming the active site cleft *(Murakami & Darst, 2003)*. Binding of sigma to the core enzyme forms the holoenzyme. In bacterial transcription the sigma subunit is essential for initiation, because of its role in promoter recognition and the formation of the transcription bubble. A key step in promoter opening is the recognition of its highly conserved -10 region by sigma 70 *(Feklistov and Darst 2011).* The sigma subunit contains 4 conserved regions, each which interact with different areas of the promoter. Sigma binds to the -10 and -35 regions of the promoter through two of its conserved regions, region 2 and region 4 respectively *(Feklistov & Darst, 2011).* Sigma region 2 contains several conserved basic and aromatic residues that interact specifically with the -10 element of the non- transcribed strand of the promoter DNA. The recognition of the -10 region occurs simultaneously with the process of strand opening.

The regions 2 and 4 of sigma 70 also interact with the coiled coil element of the RNAP β' subunit and the RNAP βsubunit flap domain respectively *(Basu et al., 2014).* Regions 2 and 4 are joined by the conserved sigma region 3.2, which occupies the RNA exit channel of RNAP, in effect blocking it. (*Kulbachinskiy & Mustaev, 2006)*. The nascent RNA must either displace sigma allowing RNAP to escape the promoter, or dissociate from the transcription complex as an abortive transcript. Eukaryotic transcription factor TFIIB maintains a similar position at the RNA exit channel of RNAP II and is proposed to interact with the nascent RNA in the same manner. This suggests a common mechanism for abortive transcription and promoter

escape in prokaryotes and eukaryotes (*Kulbachinskiy & Mustaev, 2006)*.

To aid in determining the role of sigma 3.2 in transcription initiation deletion and substitution mutants were made and used in *in vitro* transcription assays. The deletion mutants demonstrated decreased transcription in the presence of low nucleotide concentrations. (Pupov, Kuzin, Bass, & Kulbachinskiy, 2014) Yet in the presence of highnucleotide concentrations the activity of the mutant RNAPs was compensated for and full length RNAs were synthesized. *(Pupov et al., 2014)*

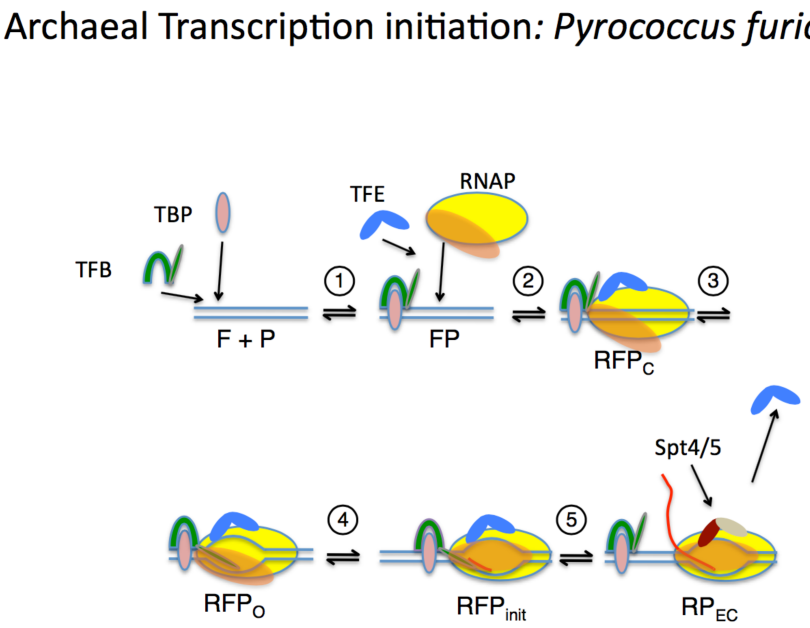



Figure 3: The transcription cycle in *P. furiosus*. (F = factors, P =promoter, R= RNAP, = open, $P_{init}$ =initiation, $P_{EC}$ = elongation complex). Sequential binding of general transcription factors marks the formation of the closed complex ($P_c$). The addition of TFE (step 2) begins the melting of the promoter to form the open complex ($P_o$). Abortive transcription ensues. When the RNA reaches a certain length the RNA polymerase enters the elongation phase Spt4/5 outcompetes TFE for binding to the RNAP clamp coiled- coil.

## 1.2 Transcription factor IIE in Eukarya

Mammalian TFIIE consists of two conserved subunits: TFIIEα and TFIIEβ, which correspond to counterparts Tfa1 and Tfa2 found in yeast *(Ohkuma & Roeder,1994).* Tfa1 contains of an N-terminal winged helix domain (WH) followed by a zinc ribbon domain (ZR). Both of these regions are essential for function of the protein *(Ohkuma, Hashimoto, Wang, Horikoshi, & Roeder, 1995).* Winged helix domains are a subclass of the helix-turn-helix (HTH) motif and consist of a fold comprised of three α-helices and three β-strands in a canonical order: α1-β1-α2-α3-β2-β3. The WH domain participates in establishing protein-DNA contacts particularly in transcription initiation complexes. WH domains are involved in DNA strand separation as in the case of DNA topoisomerase II *(Charoensawan, Wilson, & Teichmann, 2010).* Disruption of either the Tfa1 WH domain or the ZR domain in yeast proved to be a lethal phenotype, while cells with deletions in the Tfa1 C-terminus showed no ill effects and grew normally *(Grunberg et al., 2012).* TFIIEα/Tfa2 contains a central tandem WH domain. The N-terminal WH1has been shown to bind double stranded DNA *in vitro (Okuda et al., 2008) (Tanaka 2009)* but deletion of WH1 within the Tfa2 tandem WH domain caused no deleterious effects and the growth phenotype was normal. Deletion of WH2 caused a slow growth phenotype. Deletion of both WH domains was lethal. This implies a redundancy of function of the two WH motifs *(Grunberg 2012).*

The positioning of TFIIE has been studied through site-specific protein-protein cleavage experiments. Bromoacetamidobenzyl-EDTA (BABE) is a chelate labeling reagent that conjugates with sulfhydryl groups. Its iron chelate (FeBABE) is a tool with

which protein-protein, or protein-DNA interactions can be determined. The target protein is labeled on either its C or N-terminus with a substance that can be identified by an antibody. The other protein, the cutting protein, is conjugated with Fe-BABE, an EDTA chelated iron atom linked to a sulfhydral-reactive moiety. If the Fe-EDTA portion of the Fe-BABE cutting protein is located close to the target site, the protein backbone will be cleaved there defining a point of interaction. The resulting products are analyzed by gel electrophoresis and Western blot (*http://www.funakoshi.co.jp/data/datasheet/PCC/20332.pdf*). Fe-BABE cleavage experiments along with molecular modeling demonstrate the interaction of the Tfa1 WH with the RNA polymerase clamp domain, specifically at the coiled-coil region. It also participates in dimerization with the Tfa2 tandem WH domain, which spans the pol II cleft. This allows for the tandem WH domain to interact with the upstream DNA by encircling it and enabling interaction with single stranded DNA in the open complex. This places TFIIE near initial transcription bubble formation where it could potentially assist in promoter opening in addition to interaction with the non-template strand of DNA after the open complex is formed *(Grunberg 2012)*. To further elucidate the function of the WH domain in TFIIE$, mutants of the WH were made in conserved regions and used in transcription assays. These mutants demonstrated two types of deficiencies: transcription initiation and transition to elongation (Tanaka, Akimoto, Kobayashi, Hisatake, & Hanaoka, 2014).

## 1.3 Transcription factor E in archaea

Transcription factor E (TFE) found in archaea is homologous to the N-terminal portion of the eukaryotic TFIIE Tfa1 subunit (Fig.2). It has been determined that the presence of TFE is not required for transcription *in vitro*, yet all archaeal genomes encode a gene for TFE, and attempts to delete TFE in *Thermococus kodakarensis* have resultedin a lethal phenotype *(Santangelo and Reeve unpublished data).* TFE enhances formation of the transcription bubble through possible interaction with the non-transcribed (NT) strand of DNA *(Grohmann,Chakraborty et al., 2011).* The crystal structure of TFE from *Sulfolobus solfataricus* demonstrates that the N-terminus of TFE adopts an extended winged helix *(Meinhart et al., 2003)*. The C-terminus adopts a zinc ribbon (ZR) domain. (Fig. 2) The winged helix motif is commonly found in transcription factors and other DNA/RNA binding proteins. Sequence alignments demonstrate a high degree of conservation of across species (Fig. 3). It has been demonstrated that TFE is not required for transcription *in vitro*, but can stimulate transcription in the circumstance of diminished TBP-promoter recognition and in the circumstance of low TBP concentration *(Bell, Magill, & Jackson, 2001).*

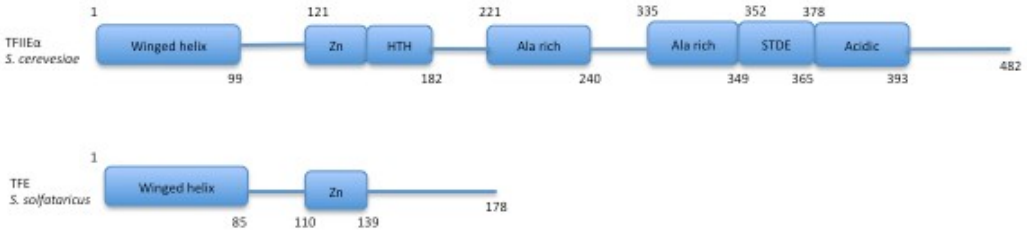



Figure 4: Primary structure comparison of the eukaryotic TFIIE$ with domain boundary numbering based on *Saccharomyces cerevisiae,* and archaeal TFE with domain boundary numbering based on *Sulfolobus solfataricus* (*Meinhart, A., Blobel, J., & Cramer, P. 2003).*

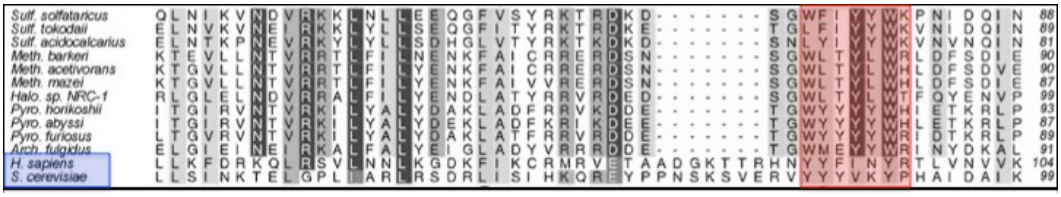



Figure 5: Sequence alignments of TFE and TFIIE. These alignments demonstrate conserved and highly conserved amino acids. Black shading represents areas of high conservation. Of special note is the aromatic patch found in the winged helix at position 76-81 (highlighted in red) in *Pyrococcus furiosus.* Figure adapted from Meinhart, et al 2003.

Photochemical crosslinking experiments were used to determine where archaeal TFE binds to DNA. A photoreactive chemical group was inserted into TFE or DNA and upon exposure to UV light a crosslink formed between DNA and protein in close proximity. It was demonstrated previously that TFE binds specifically to the non- transcribed (NT) strand of DNA in the transcription bubble at positions -9 and -11 *(Grunberg, Bartlett, Naji, & Thomm, 2007).* Current cross-linking studies indicate that the tip of the TFE winged helix is close to the DNA

(*Brown and Bartlett, unbpublished results)*.

To investigate the positioning of TFE in the archaeal transcription initiation complexes single molecule FRET (Forster Resonance Energy Transfer) experiments were used. smFRET requires the binding of fluorophores to two distinct molecules. Should they come in proximity to one another energy can be transferred from the donor fluorophore to the acceptor. This causes an excitation of the fluorophore that is detected through fluorescent microscopy. smFRET experiments localized the interaction of the TFE winged helix and the elongation factor Spt 4/5 to the same region within the RNA polymerase clamp. TFE from *Methanococcus jannaschii* was labeled on the WH, and the ZR with a fluorescent probe. Complementary fluorescent probes were incorporated into five reference sites on the RNAP, PICs were formed, the complexes were immobilized and examined, and FRET efficiencies were determined. It was determined that the WH domain binds within the RNAP clamp near the coiled coil and the ZR domain binds to an area between the stalk and the clamp *(Grohmann et al., 2011).*

Previous work had determined the binding site of Spt4/5 corresponds to the coiled-coil region within the clamp, the same area that TFE WH binds. Binding competition experiments were performed and compared the effects of TFE and Spt4/5 on initiation and elongation phases of transcription. It was discovered that during initiation Spt4/5 inhibits transcription but TFE can displace Spt4/5 and overcome this inhibition. During elongation Spt4/5 displaces TFE so that elongation can effectively occur *(Grohmann et al., 2011).*

**1.5 Current understanding of TFE structure and function.**

Sequence alignments of TFE demonstrate a high degree of conservation between archaea and the N-terminus of the $ subunit of TFIIE in eukarya. (Figs. 2 and 3). The N terminus of TFIIE and TFE both contain a winged helix (WH) domain. WH domains are found in transcription systems of both eukaryotes and prokaryotes *(Teichmann, Dumay- Odelot, & Fribourg, 2012)*. The winged-helix (WH) domain of TFIIE$ and TFE contains a well-conserved aromatic patch near the tip of the wing. Aromatic regions of proteins have been shown to be involved in opening DNA through base stacking interactions with nucleotides. A conserved aromatic patch in sigma region 2.3 in *Escherichia coli* plays a role in promoter opening through specific base stacking interactions near the -10 element *(Feklistov & Darst, 2011)*. Recent structural comparisons between sigma, TFB, and TFIIB imply an evolutionary link between the three proteins *(Taylor, Burton, Burton, 2014).*

The archaeon *Pyrococcus furiosus* contains two TFIIB homologs, TFB1 and TFB. *In vitro* transcription assays using TFB2 have demonstrated reduced transcription as compared to transcription with TFB1 *(Micorescu et al., 2008).* TFB2 is a variant of TFB1 that lacks the B-finger region, which is important for transcription start site selection.

Experimentation has shown the presence of TFE in transcription assays with TFB2 can partially alleviate this defect in transcription (*Micorescu et al., 2008*). In eukaryotes during transcription initiation TFIIE$ binds to Rpb 1 and Rpb2 which are the largest and second largest subunits respectively, of the RNA polymerase *(Tanaka et al., 2014).*

These subunits make up the clamp and jaw of the polymerase and contain the active site of transcription *(Hahn, 2005)*. Recent experiments have shown that mutations in the winged helix domain of TFIIE$ affect two important components of transcription: initiation and the transition from initiation to elongation *(Tanaka et al., 2014)*. Mutations in the C- terminal end of the WH in TFIIE that correspond to the aromatic patch in TFE demonstrate a decrease in transcription in the presence of negatively supercoiled DNA and linear template. These same mutations demonstrate a diminished ability to bind to TFIIE *(Tanaka et al., 2014)*.

The experiments in this thesis address the role of specific sequences of TFE in *P. furiosus* in transcription activation and interactions with TFB.

# CHAPTER 2

## RESULTS

TFE, which is encoded in all sequenced archaeal genomes, is homologous to the N-terminal domain of the alpha subunit of TFIIE in eukarya. TFE consists of a winged helix (WH) domain and an N-terminal zinc ribbon. Near the C-terminus of the TFE/TFIIE winged helix is a patch of aromatic amino acids (Figures 3 and 4). This patch of aromatic amino acids is highly conserved in archaea and eukaryotes, implying an important role in transcription. Aromatic amino acids are known to interact with nucleotides in base stacking interactions as in the case of transcription bubble formation. This is demonstrated in the bacterial system by sigma region 2.3; this region contains a number of invariant aromatic amino acids. These aromatic amino acids interact specifically with the -10 site upstream of the promoter and nucleate transcription bubble formation *(Feklistov & Darst, 2011)*.

*Pyrococcus furiosus* encodes two TFB paralogs, TFB1 and TFB2 that are homologous to the eukaryotic transcription factor TFIIB. TFB2 diverges from TFB1 in that it lacks parts of the regions known as the B-reader and B-linker, are involved in promoter opening and transcription start site selection *(Micorescu et al., 2008)*. Transcription with TFB2 *in vitro* is deficient. It has been demonstrated that the presence of TFE can facilitate transcription in cases where TFB function is suboptimal as in the case of TFB2 *(Micorescu et al., 2008)*.

I predict that the aromatic patch in the TFE winged helix domain plays a key role in the initiation of transcription in the presence of the deficient TFB protein, TFB2.

In *Pyrococcus furiosus,* the TFE aromatic patch consists of four tyrosines flanked on either side by a tryptophan. Previous photochemical crosslinking experiments demonstrated an interaction with some of these aromatic amino acids and nearby amino acids to the non- transcribed strand of DNA (*Brown and Bartlett, 2013, unpublished data*).

To further elucidate the function of this aromatic patch and other amino acids in proximity, several mutations were made and the mutant proteins were tested for function. These mutations were as follows: W76, Y77, Y78, Y79, Y80, W81 (Figure 4) were all substituted to alanines as a whole and individually. Also, two triplet mutations were made: one for the first half of the aromatic patch, W76, Y77, Y78 and the other of the second half Y79, Y80, W81, in which all were substituted with alanines. In addition, R70, and F66, which lie in proximity to the aromatic patch and have demonstrated crosslinking to the NT strand, were also mutated to alanines. These TFE mutants were transformed into BL21gold *Escherichia coli* cells and the proteins were overexpressed.

The mutant proteins were subsequently purified on a His-Pur™ cobalt resin column using a native preparation protocol (see materials and methods).

To test TFE for function, mutants were used in transcription assays with the strong, well-characterized, glutamate dehydrogenase promoter (gdhP), purified RNA polymerase, TBP, and TFB1 or TFB2. Due to the minimal effect of TFE on transcription *in vitro* in the presence of TFB1, TFB2 was used to aid in determining the function of the patch as it has previously demonstrated response to the presence of TFE.

Figure 6A.

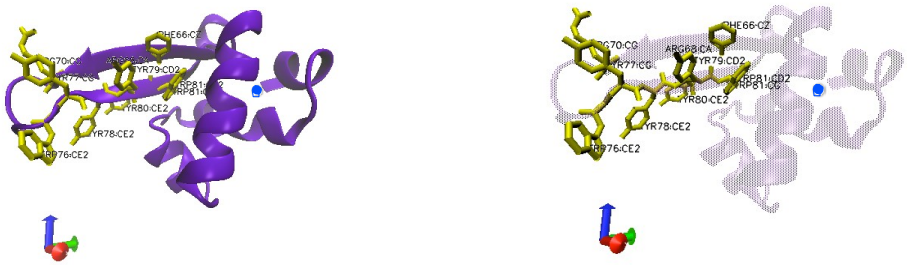



Figure 6B.

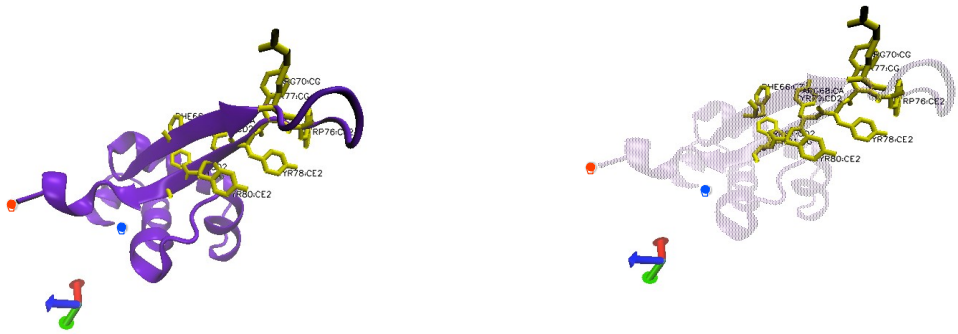



Figure 6: A. Structural model for *P. furious* TFE. *Pfu* TFE in cartoon representation with the amino acids mutated to alanines represented as wireframe. Model created from *Pyrococcus furious* TFE sequence using Swiss Model homology builder *(Kiefer, Arnold, Künzli, Bordoli, & Schwede, 2009)*. Images generated using VMD 1.9.2 software. Red dot indicates C terminus, Blue dot indicates N- terminus. B. Same as A but rotated 180°. The views on the right are transparent cartoon view to enhance resolution of the mutated amino acids *(Kiefer, Arnold, Künzli, Bordoli, & Schwede, 2009)*.

## 2.1 Standard transcription assays with mutants

To define the effects of mutations in the TFE aromatic patch on transcription initiation, *in vitro* transcription reactions were performed with either TFB1 or TFB2 in the presence of specific TFE mutants. Figure 7A shows the activity of the TFE mutants in the presence of TFB1 and TFB2. Lane 1 shows transcription with DNA only, and as expected no transcript is observed. Lane 2 shows transcription with DNA and RNAP only and as expected no transcript is observed. Lanes 3 and 4 show transcription with the

addition of TFB1 and TFB2 respectively. Lane 4, with TFB2 shows a decreased level of transcription as compared to lane 3 with TFB1. The addition of WT TFE shows no effect with TFB1 (compare lanes 3 and 5) but a moderate increase in the amount of run-off transcript with TFB2 (compare lanes 4 and 6). Lanes 7 and 8 contain TFB1 and TFB2 respectively in the presence of the 6 alanine substituted aromatic patch in TFE. Transcription with TFB1 is unaffected but with TFB2 the transcription efficiency is similar to the absence of TFE (lane 4). Transcription efficiencies of the alanine substituted point mutants with TFB1 (lanes 9,11,13) appear similar to lane 5 and those with TFB2 (10,12,14) appear similar to lane 6 that contains WT TFE.

Figure 7A.

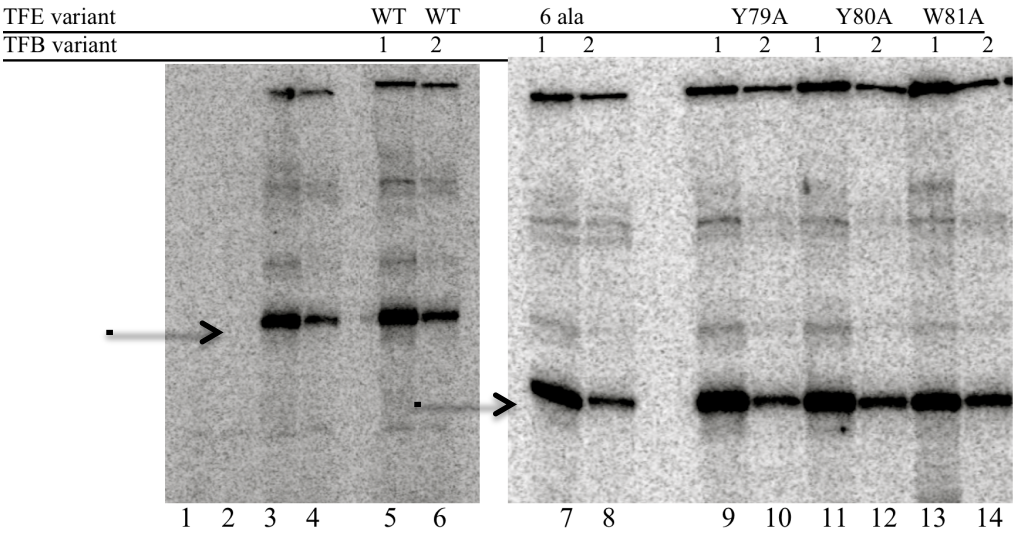



Figure 7A: Winged helix mutants and their effect on transcription. TFB1 or TFB2 were used in transcription assays with the gdh promoter -60 to +37 in the presence of a specific mutant. Lanes 1-8 in each gel image represent the controls. The arrow indicates the location of run off transcript. The TFB variant and specific mutant is indicated at the top of each lane. Transcription assays were performed as indicated in materials and methods. After transcription the radioactively labeled reactions were subjected to urea-PAGE and detected by autoradiography. Radioactivity at the top of the gel represents material caught in the wells. Higher molecular weight bands likely arise from template switching of transcribing RNA polymerase. C-terminal winged helix point mutants and their effect on transcription.

Figure 7B shows the comparison of the mutants in and near the N-terminus of the winged helix on transcription efficiencies in the presence of TFB1 and TFB2. Lanes 1 and 2 demonstrate no transcription in the absence of RNAP and TFB (lane 1) and TFB (lane 2). Lanes 3 and 4 show transcription with the addition of TFB1 (lane 3) and TFB2 (lane 4). Lanes 5 and 6 show transcription of TFB1 (lane 5) and TFB2 (lane 6) with the addition of WT TFE. Once again the presence of WT TFE shows a moderate increase in transcription in the sample with TFB2 as compared to TFB2 alone. TFB1 in the

presence of TFE mutants (lanes 9,11,13,15,17) demonstrate transcription efficiencies similar to that of lane 5. TFB2 in the presence of TFE mutants (lanes 10,12,14,16,18) demonstrate transcription efficiencies similar to lane 6.

Figure 7B.

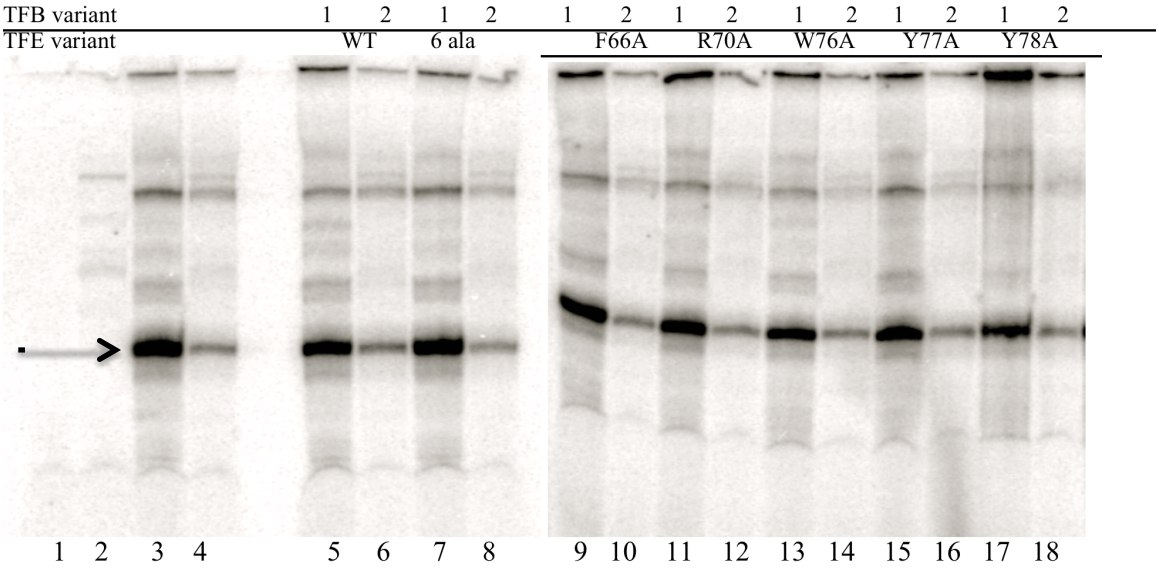



Figure 7B: Winged helix mutants and their effect on transcription. TFB1 or TFB2 were used in transcription assays with the gdh promoter -60 to +37 in the presence of a specific mutant. Lanes 1-8 in each gel image represent the controls. The arrow indicates the location of run off transcript. The TFB variant and specific mutant is indicated at the top of each lane. Transcription assays were performed as indicated in materials and methods. After transcription the radioactively labeled reactions were subjected to urea-PAGE and detected by autoradiography. Radioactivity at the top of the gel represents material caught in the wells. Higher molecular weight bands likely arise from template switching of transcribing RNA polymerase.

Transcription assays using the triplet mutations were performed as previously described and the results are demonstrated in 7C. Lanes 1 and 2 demonstrate

no transcription in the absence of RNAP and TFB (lane 1) and TFB (lane 2). Lanes 3 and 4 show transcription with the addition of TFB1 (lane 3) and TFB2 (lane 4). Lanes 5 and 6 show transcription of TFB1 (lane 5) and TFB2 (lane 6) with the addition of WT TFE. There is a moderate increase in transcription in the lane with TFB2 and the addition of TFE (compare lanes 4 and 6). Lanes 7 and 8 contain TFB1 and TFB2 respectively both with the addition of the aromatic patch 6 alanine substitution. Lane 7 containing TFB1 appears unaffected by the mutant but TFB2 transcription levels appear similar to lane 4 with TFB2 only. The N-terminal triplet mutant, 76-WYY -78 to AAA, in lanes 9 and 10 appear similar in transcription efficiencies to that of lanes 5 and 6, which contain WT TFE. Lanes 11 and 12 contain the C-terminal triplet mutant, 79-YYW-81 to AAA, also have transcription efficiencies that are similar to those of lanes 5 and 6 respectively.

Figure 7C.

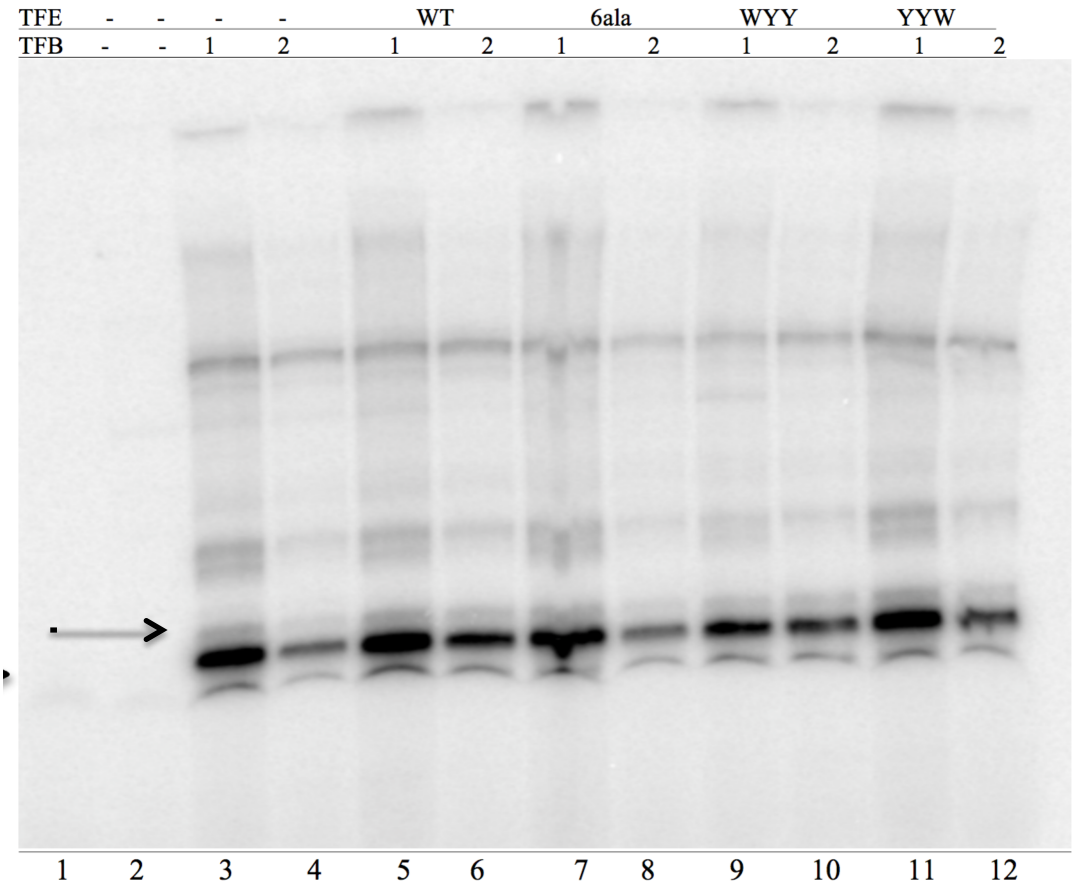
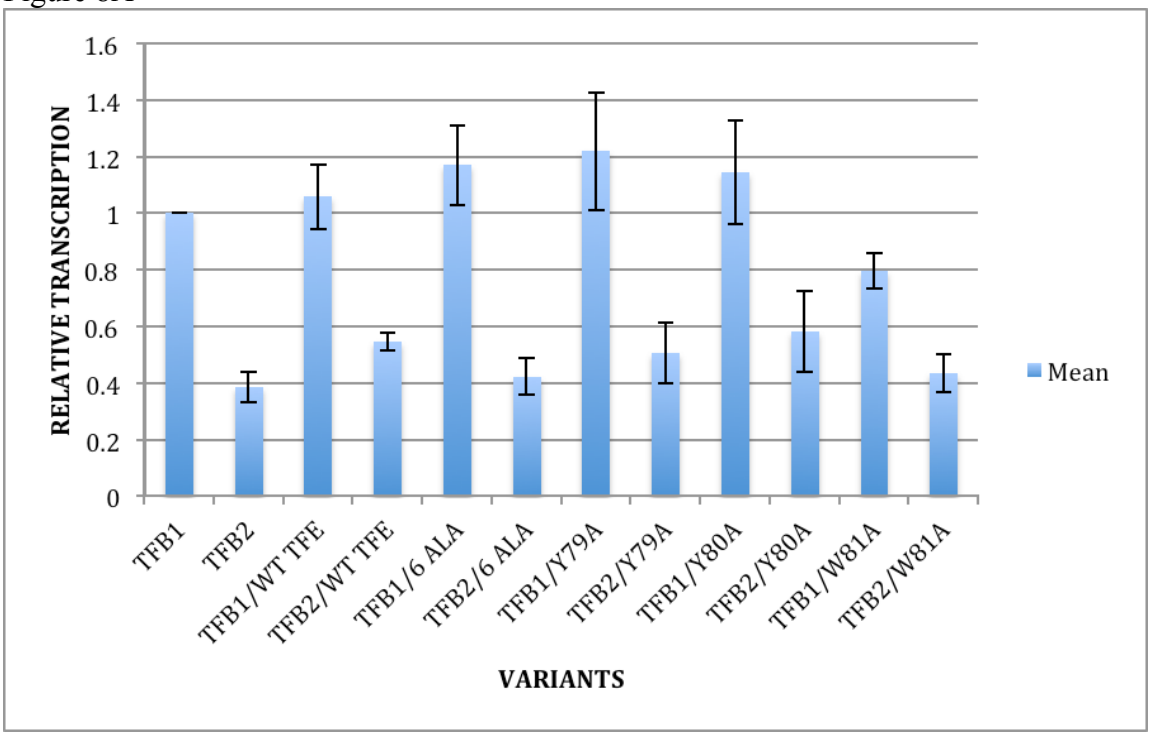



Figure 7C: Winged helix mutants and their effect on transcription. TFB1 or TFB2 were used in transcription assays with the gdh promoter -60 to +37 in the presence of a specific mutant. Lanes 1-8 in each gel image represent the controls. The arrow indicates the location of run off transcript. The TFB variant and specific mutant is indicated at the top of each lane. Transcription assays were performed as indicated in materials and methods. After transcription the radioactively labeled reactions were subjected to urea-PAGE and detected by autoradiography. Radioactivity at the top of the gel represents material caught in the wells. Higher molecular weight bands likely arise from template switching of transcribing RNA polymerase. C-terminal winged helix point mutants and their effect on transcription.

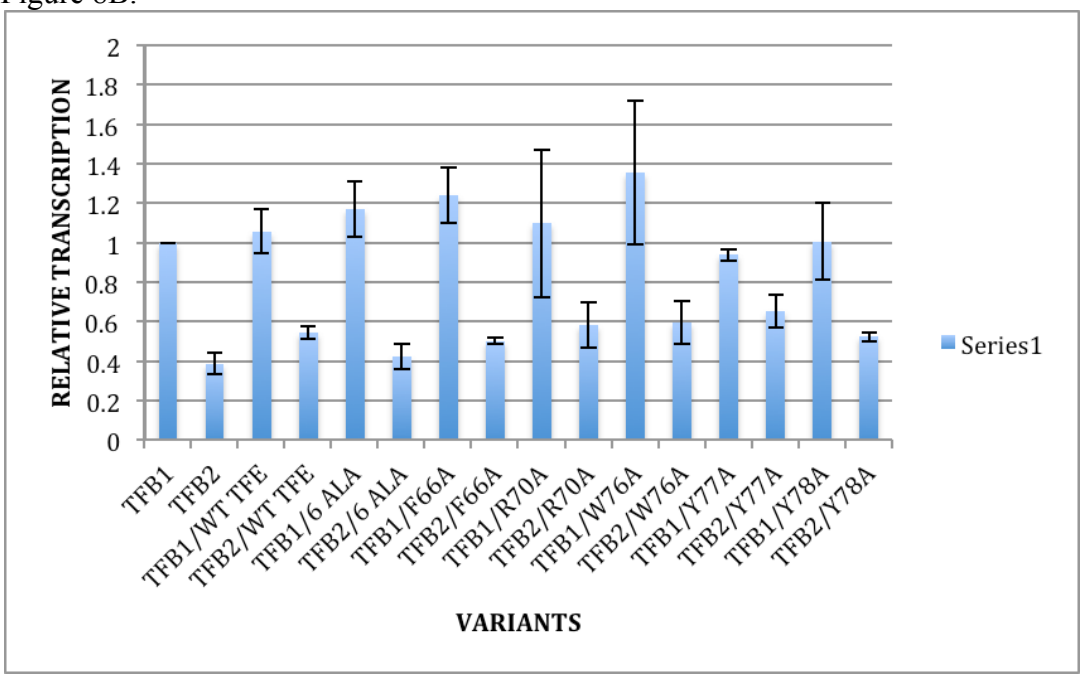
Figure 8 illustrates quantitation of transcription efficiencies of the C-terminal point mutations in the presence of TFB1 and TFB2 as compared to WT TFE and the 6 alanine substituted TFE. Transcription efficiencies of TFB2 are consistently lower than that of TFB1, even in the presence of WT TFE. There is a moderate increase in transcription in the presence of WT TFE, but none with the TFE 6 alanine mutant while the individual substitutions appear to have little or no effect on activation of TFB2 in this assay.

Figure 8A




8A. Histogram of the relative transcription production of three experiments using the C-terminal TFE mutants normalized to TFB1 transcription rates (see figure 5A). Error bars indicate standard error. Relative transcription efficiencies were calculated using ImageQuantTL software. Band pixelation density was determined by drawing a box around the band and a numerical value was given. Background was subtracted by volume measurement of an area equivalent to the measured band and subtracted from the numerical value. This value was averaged between all the samples of the same type. TFB1 values were set to 1 and each data type was given a value based on its relative density in comparison to TFB1.

Figure 8B.




8B. Histogram showing relative transcription rate averages of three experiments using single point mutants in or around the winged helix. (see figure 5B). Error bars indicate standard error. Relative transcription efficiencies were calculated using ImageQuantTL software. Band pixelation density was determined by drawing a box around the band and a numerical value was given. Background was subtracted by volume measurement of an area equivalent to the measured band and subtracted from the numerical value. This value was averaged between all the samples of the same type. TFB1 values were set to 1 and each data type was given a value based on its relative density in comparison to TFB1.

Figure 8C.

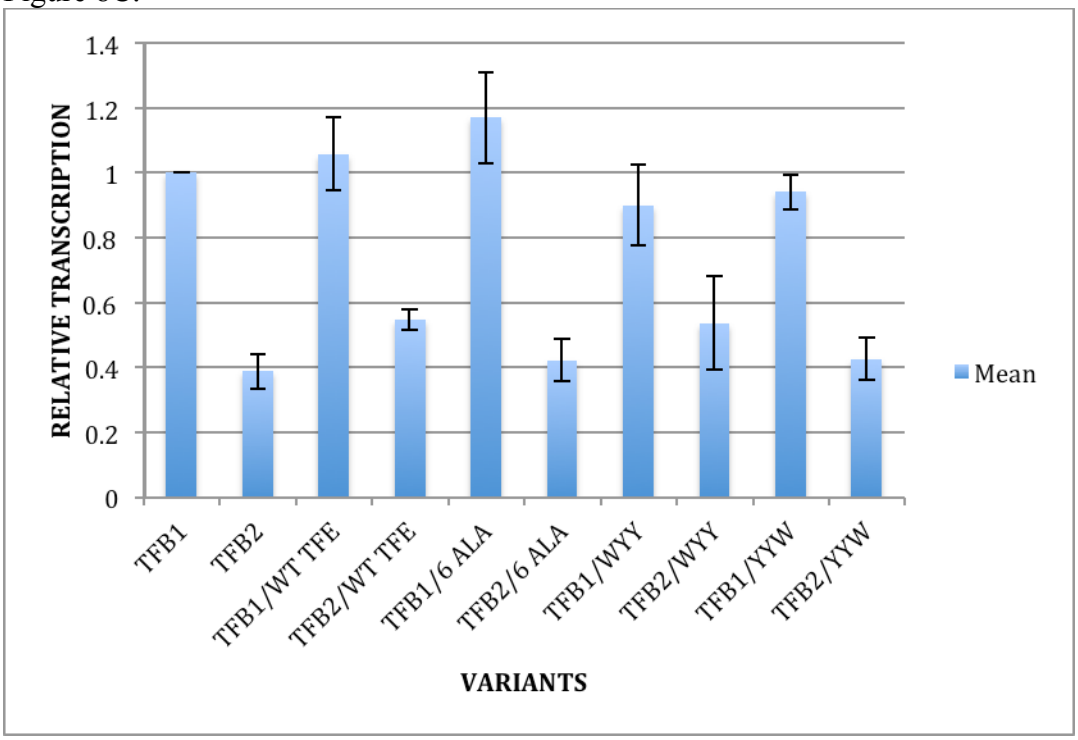



Figure 8C: Histogram showing relative transcription efficiencies of the triplet mutations (see figure 5C). Error bars indicate standard error. Relative transcription efficiencies were calculated using ImageQuantTL software. Band pixelation density was determined by drawing a box around the band and a numerical value was given. Background was subtracted by volume measurement of an area equivalent to the measured band and subtracted from the numerical value. This value was averaged between all the samples of the same type. TFB1 values were set to 1 and each data type was given a value based on its relative density in comparison to TFB1.

**2.2 Transcription assays with heparin challenge.**

The stability of protein-DNA complexes can be tested by challenge with the polyanion, heparin. Heparin is a glycosaminoglycan and is used medically as an anticoagulant *(Nelson & Cox, 2005).* Heparin is highly sulfated and has the highest negative charge density of any known biological molecule. *(Nelson & Cox, 2005).* For this reason it is often used as a general competitor of DNA in transcription assays, since it mimics the negatively charged DNA phosphate sugar backbone. Any DNA binding protein, such as RNA polymerase, with an affinity for a negatively charged polymer will interact with excess heparin and be prevented from binding to DNA. I predicted that preinitation complexes formed with TFB2 might not be as stable as those formed with TFB1 and, therefore, more susceptible to increased concentrations of heparin. I also predicted that TFE could assist in stabilizing the TFB2-containing PIC in the presence of an increased concentration of heparin and that this could provide a better assay for determining TFE function.

To test the effects of heparin on complexes transcription assays were performed using the gdh promoter with purified proteins. Increasing concentrations of heparin were added after the initial pre-incubation of RNAP, TBP, and TFB1 or TFB2 (see figure 9). Lanes 1 and 2 contain DNA only and DNA and RNAP respectively.Transcription with TFB2 in the presence of increased concentration of heparin is diminished (lanes 5, 6, 7) as compared to lane 4 with 50μg/mL.

Figure 9:

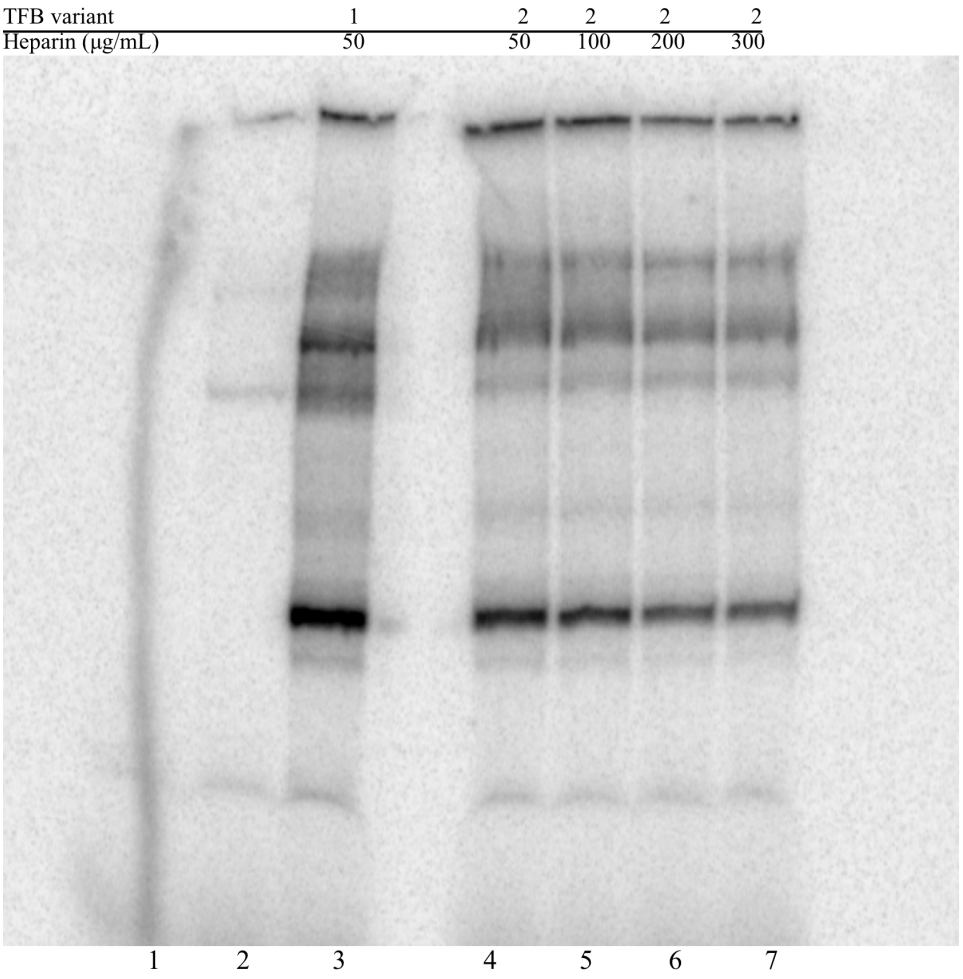



Figure 9: The effects of increasing concentration of heparin on transcription complexes formed with TFB2. Transcription complexes were formed using the gdh promoter -60-+37. TFB variant is indicated above each lane. After transcription the radioactively labeled reactions were subjected to urea-PAGE and detected by autoradiography. The concentration of heparin is also indicated above each lane, and was added to pre-initiation complexes 30 seconds before NTPs were added to allow transcription. The 37 nucleotide transcript is indicated by the arrow.

To further investigate the influence of heparin on transcription, a higher concentration of heparin at a constant of 300 µg/mL was used and the incubation times were varied

(Figure 10). Lanes 1 and 2 are DNA only and DNA plus RNAP respectively. Each of those samples was also incubated with heparin for 30s. No transcript was observed, as was expected. The resulting transcription efficiencies of lanes 3-6 with TFB2 appear similar. This suggests that the transcription complexes do not dissociate significantly in this time frame.

Figure 10:

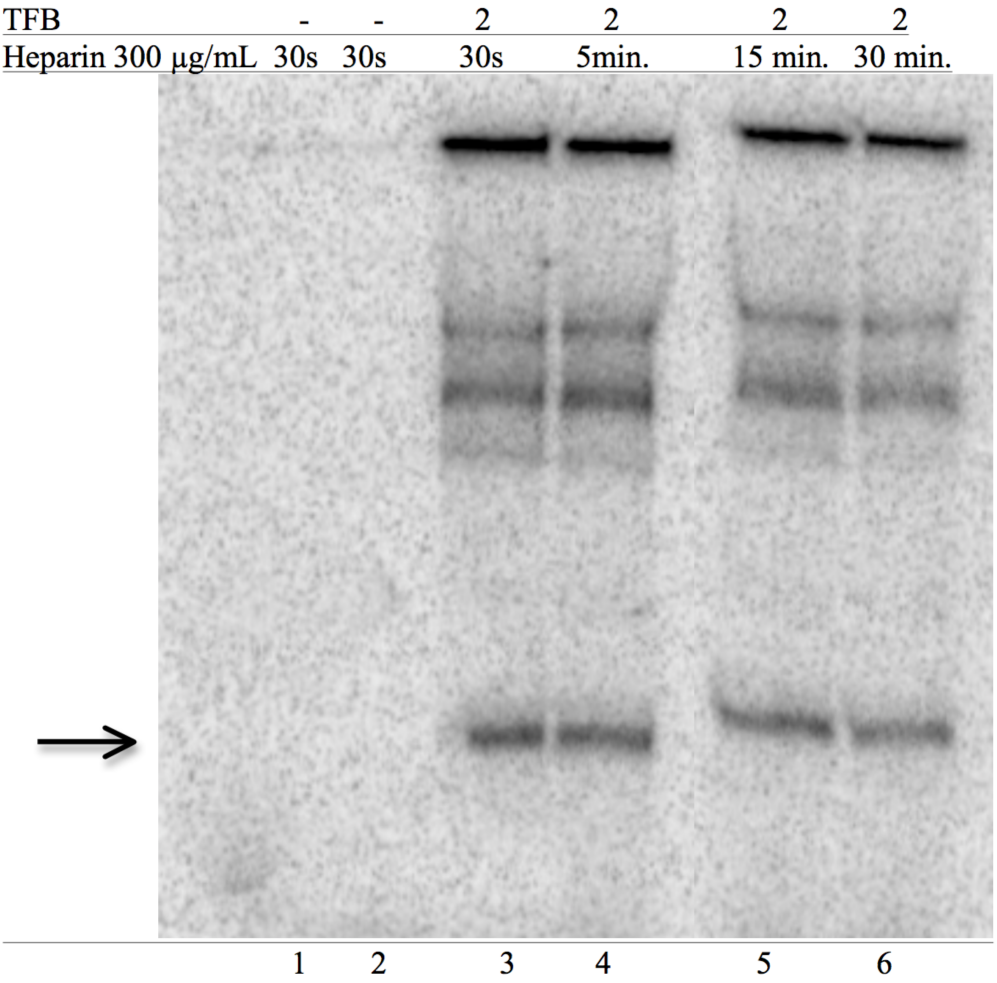



Figure 10: Transcription complexes Effect of increased time of heparin challenge on transcription. Transcription complexes were formed using the gdh promoter -60-+37. After transcription the radioactively labeled reactions were subjected to urea-PAGE and detected by autoradiography. Lane 1 contains DNA only and lane 2 contains RNAP and DNA only. TFB variant is indicated above each lane. The concentration of heparin is also indicated above each lane Increasing lengths of time of exposure to this concentration of heparin were performed from 5 minutes to 30 minutes. Arrow represents run-off transcript.

To test whether TFE can increase transcription in the presence of heparin, TFE was added prior to the addition of heparin (Figure 11). Lane 1 shows transcription with DNA only, and as expected no transcript is observed. Lane 2 shows transcription with DNA and RNAP only and as expected no transcript is observed. Lane 3 contains TFB1 and lane 4 contains TFB2. Lanes 1-4 were challenged with 50 µg/mL heparin and serve as controls for this experiment. Lane 5 contains TFB2 and is challenged with 300 µg/mL of heparin without TFE. Lane 6 contains TFB2 and TFE and is challenged with 50 µg/mL heparin. Lane 7 contains TFB2 and TFE challenged with 300 µg/mL of heparin. (Compare transcription efficiencies of lanes 4 to lane 6 and lane 5 to lane 7). There is a mild increase in transcription in the presence of high concentrations of heparin with the addition of TFE.

Figure 11:

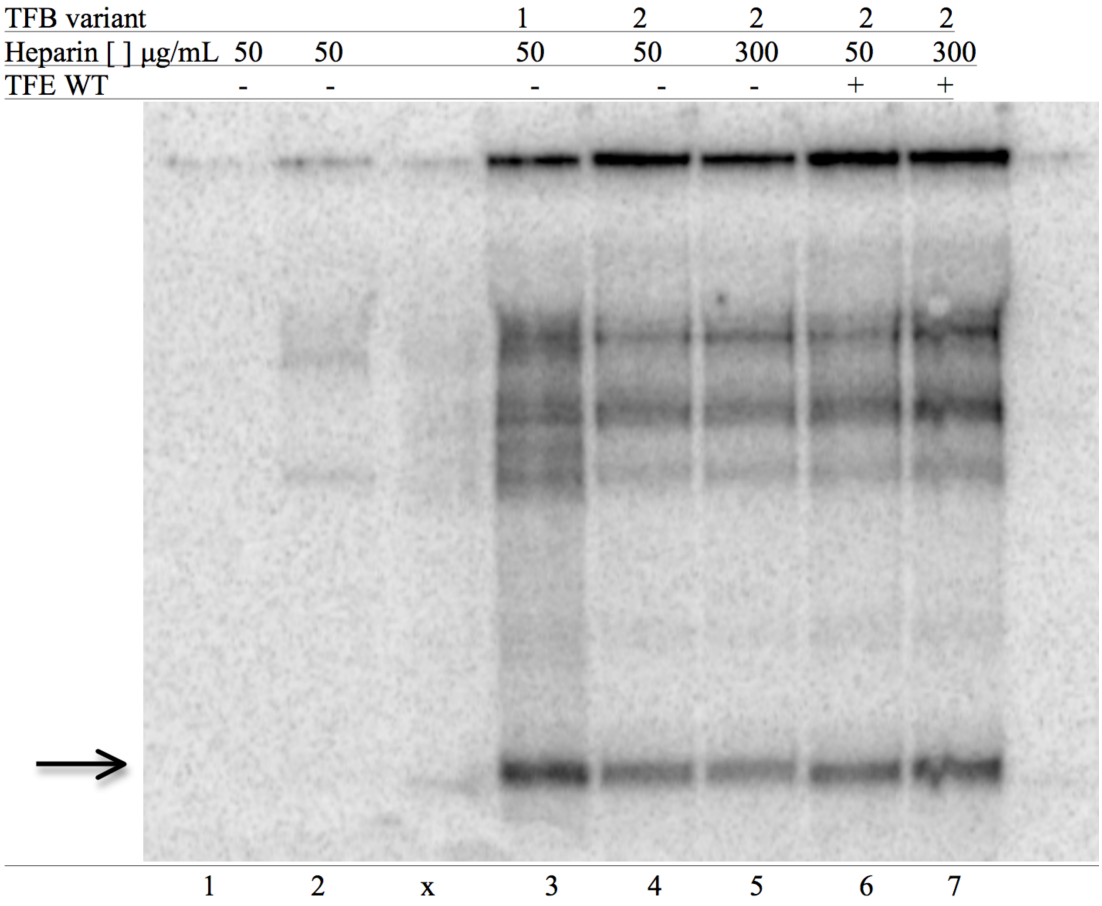



Figure 11: Effect of the addition of TFE to heparin challenged complexes. Transcription complexes were formed using the gdh promoter -60-+37. TFB variant is indicated above each lane. The concentration of heparin is also indicated above each lane (Banding in lane x represents overflow from lane 3). After transcription the radioactively labeled reactions were subjected to urea-PAGE and detected by autoradiography. The 37 nucleotide transcript is indicated by the arrow.

To test the idea that a high concentration of heparin may have an even greater effect on unstable PICs, and that TFE might reverse this effect, the heparin concentration was increased to 1,000ug/mL and transcription was measured. To evaluate the contribution of the aromatic patch to the stability of the complexes, both WT and the WYYYYW to AAAAAA substituted aromatic patch TFE variants were added. Figure 12 demonstrates that in the presence of an unmodified TFE, in the assays with TFB2, there is a slight increase in transcription with a high concentration of heparin. Lanes 1 and 2 contain DNA and DNA and RNAP only respectively. Both were challenged with 1,000%g/mL of heparin. Lanes 3 and 4 are TFB1 and TFB2 also challenged with 1,000%g/mL of heparin. Lanes 5 and 6 include the addition of WT TFE and lanes 7 and 8 show the effects of the addition of the 6 alanine substituted aromatic patch of the winged helix. While the higher heparin concentration reduced TFB2 mediated transcription, TFE was not able to rescue this low activity. This suggests that TFE does not increase the stability of transcription complexes or if so it is minimal.

Figure 12:

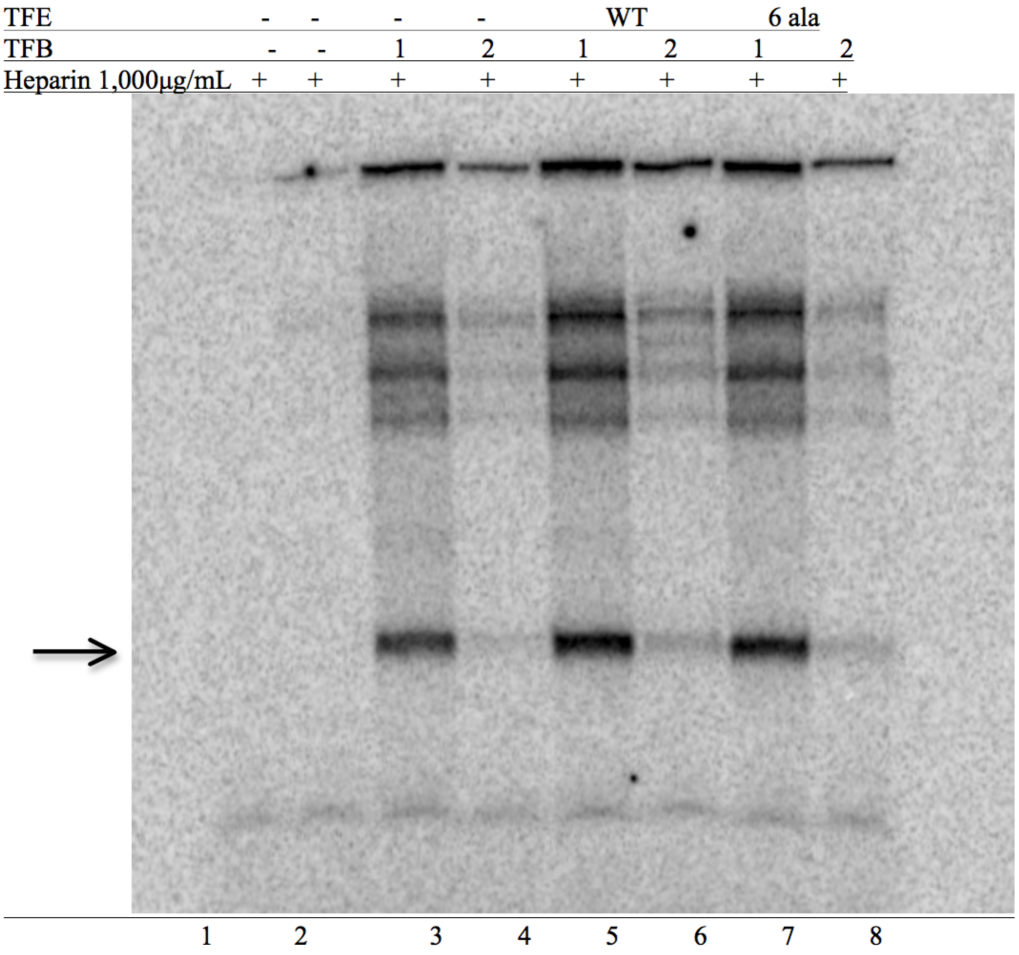



Figure 12: Transcription complexes challenged with 1,000 µg/mL heparin with the addition of WT TFE and the 6 alanine substituted aromatic patch mutant of TFE. Transcription complexes were formed using the gdh promoter -60-+37. TFB variant is indicated above each lane. The concentration of heparin is also indicated above each lane. After transcription the radioactively labeled reactions were subjected to urea-PAGE and detected by autoradiography. The 37 nucleotide transcript is indicated by the arrow.

**2.3 Transcription assays with low versus high nucleotide concentration.**

*P. furiosus* TFB2 is missing regions of the conserved B-reader, an archaeal structure analogous to region 3.2 of the bacterial sigma factor. Based on earlier experimentation by Pupov et al (*Pupov et al., 2014*), using *E. coli* sigma 70 region 3.2 in the presence of low nucleotide conditions I predict that the TFB B-reader plays roles similar to that of bacterial sigma region 3.2 in bacteria. Since TFB2 is highly divergent in this region relative to TFB1, I predict that the TFB2 transcription complex would be more sensitive to low NTP concentration than TFB1. In addition, I predict that TFE may help overcome the sensitivity of TFB2 to low NTP levels.

To test the effects of low NTP concentrations, standard transcription reactions were performed with TFB1 or TFB2. NTP levels were varied for ATP, CTP, and GTP while UTP concentrations remained low and constant, a requirement for radiolabeling (see Figure 13). Lanes 1 and 2 contain DNA only and DNA and RNAP only respectively. 500%M final concentration of adenine, cytosine and guanine along with 10%M final concentration of uracil were added after incubation with 50µg/mL of heparin. Lanes 3 and 4 contain TFB1 and TFB2 respectively with the addition of NTPs to a final concentration of 500µM/10µM (A, C, G/U). Lanes 5 and 6 have 50µM/10µM final NTP concentration to TFB1 and TFB2. Lanes 7 and 8 contain 40µM/10µM final NTP concentration. Lanes 9 and 10 contain 30µM/10µM final NTP concentration. Lanes 11 and 12 contain 20µM/10µM final NTP concentration. Lanes 13 and 14 contain 10µM/10µM final NTP concentration, and lanes 15 and 16 contain 5µM/5µM final NTP concentration.

Transcription with TFB1 is unaffected by the NTP concentration while transcription efficiencies in the presence of TFB2 are highly deficient and become increasingly so as the NTP concentration is reduced.

Figure 13:

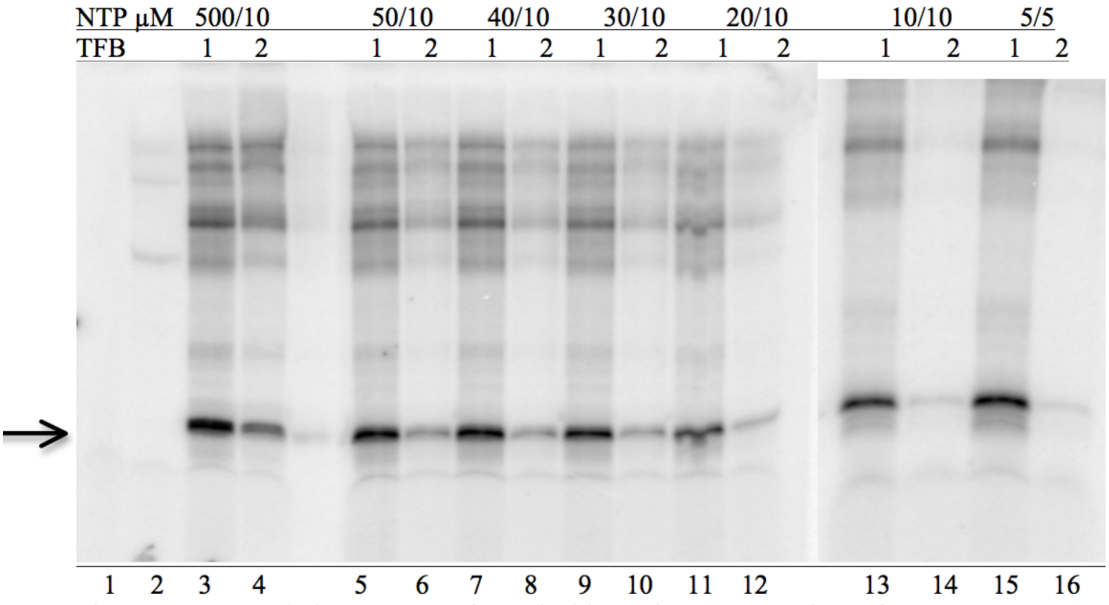



Figure 13: Transcription assays performed with varying concentrations of NTPs. Transcription using the gdh promoter -60-+37 with either TFB1 or TFB2. NTP concentrations were systematically varied with ATP, CTP, and GTP first followed by UTP and are indicated above each lane. After transcription the radioactively labeled reactions were subjected to urea-PAGE and detected by autoradiography. Numbers represent final concentration in µM. ATP, CTP, and GTP over UTP and $^{alpha32}$P- UTP. The 37 nucleotide transcript is indicated by the arrow.

To test whether TFE can rescue transcription in the presence of low nucleotide concentration. (Figure 14) Transcription complexes were formed in the presence of WT versus mutant TFE and transcription was performed with either high NTP concentration (500µM/10µM final) or low concentration (20µM/10µM final) Figure 14 demonstrates the effects of low nucleotide concentration on transcription assays with TFB1 and

TFB2. Lanes 1 and 9 contain DNA only and as expected, no transcript is observed. Lanes 2 and 10 contain DNA and RNAP only and as expected, no transcript is observed. Lanes 1-8 have a final NTP concentration of 20μM/10μM. TFB1 and TFB2 are added to the other components and the transcription efficiencies are documented in lanes 3 and 4 respectively. Lanes 5 and 6 contain TFB1 and TFB2 with the addition of TFE. The six alanine substituted TFE is added in lanes 7 and 8. Lanes 9-16 have a final NTP concentration of 500μM/10μM. Lanes 11 and 12 contain TFB1 and TFB2 respectively. Lanes 13 and 14 have the addition of TFE and lanes 15 and 16 contain the 6 alanine substituted TFE. TFB1 is less sensitive to low nucleotide concentrations, compared to TFB2, and is not affected by WT TFE. The transcription efficiency in the samples with lowered nucleotide concentrations in the presence of TFB2 show diminished transcription. The presence of TFE increases transcription in both sets of samples with TFB2. The sample with the low NTP concentration shows the greatest overall increase (Figure 14 compare lanes 4 and 6 with lanes 12 and 14). Reducing A, C, G concentrations had only a small effect on TFB1 transcription but a much larger effect on TFB2 transcription showing that TFB2 is sensitive to low NTP concentration and the presence of WT TFE increased transcription in both sets of samples with TFB2.

Figure 14:

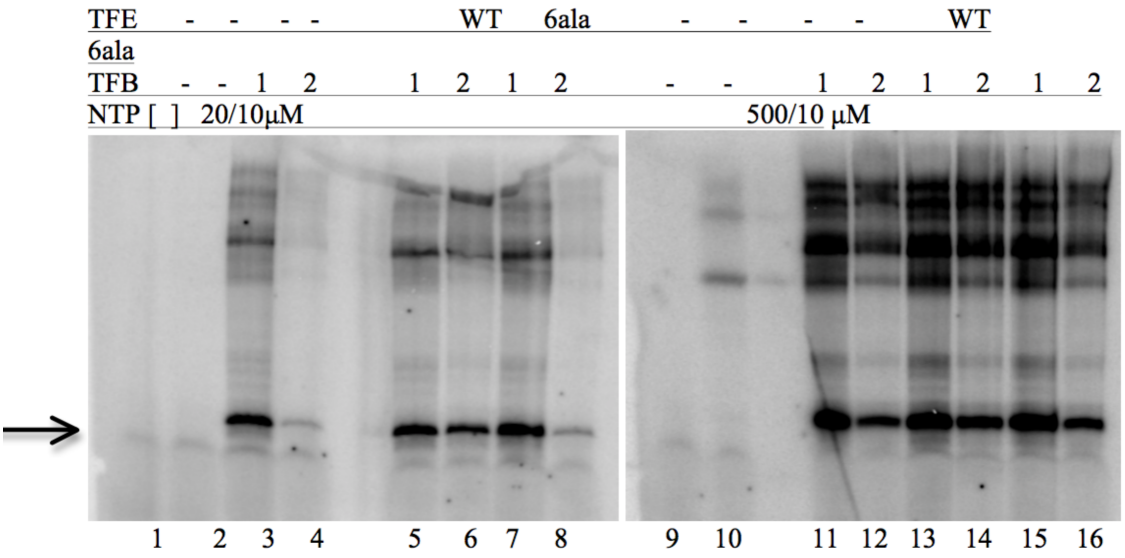



Figure 14: Effects of TFE on low nucleotide condition Transcription complexes were formed using the gdh promoter -60-+37. TFB and TFE variant is indicated above each lane. After transcription the radioactively labeled reactions were subjected to urea-PAGE and detected by autoradiography. Low vs high NTP concentration (20μM/10μM vs 500μM/10μM also indicated above each lane) comparisons made in the presence of WT and 6 alanine substituted aromatic patch TFE. The 37 nucleotide transcript is indicated by the arrow.

Figure 15:

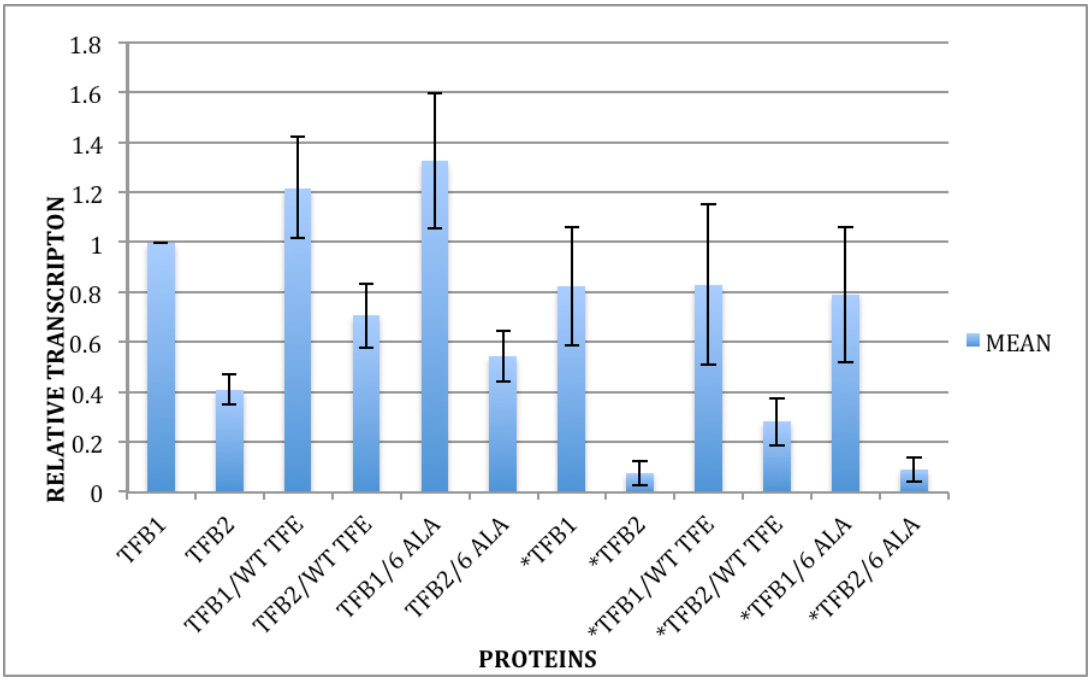



Figure 15: Comparison of transcription between high and low nucleotide concentrations. Bars represent the average of three experiments. Asterisks represent low nucleotide concentrations. Bars represent mean. All samples normalized to TFB1 transcription at high nucleotide conditions. Error bars represent standard error.

Comparison of transcription between low and high nucleotide concentration demonstrates the ability of TFE to overcome the effects of low nucleotide concentration. Assays performed with TFB1 do not appear to be adversely affected by low NTP concentrations (See figure 14). TFB1 at low nucleotide concentration is approximately 83% of that of high concentration. TFB2 transcription at high nucleotide concentration is 41% of TFB1 transcription and 7% of TFB1 transcription levels at low nucleotide concentration. Under high nucleotide conditions the addition of wild type

TFE increases TFB2 transcription by 1.7 fold. Under low nucleotide conditions the addition of wild type TFE increases transcription by 3.7 fold. When the TFE mutant is used in TFB2 transcription, only a small increase in transcription is observed.

| | | **TFE wt** | **TFE 6-ala** |
|---|---|---|---|
| **TFB1** | High NTP | 1.2 | 1.3 |
| | Low NTP | 1.0 | 1.0 |
| **TFB2** | High NTP | 1.7 | 1.3 |
| | Low NTP | 3.7 | 1.2 |

**Table 1.** Fold increase of TFB1 or TFB2 transcription by wild type or mutant TFE.

In summary, transcription with TFB2 is compromised at low NTP concentrations. Wild type TFE can compensate for low transcription with TFB2 while having little effect on TFB2 transcription at high NTP concentration, and little effect on TFB1 transcription at either high or low NTP concentration (Table 1).

# CHAPTER 3

## EXPERIMENTAL PROCEDURES

### 3.1 Mutagenesis of TFE

Mutagenesis of the entire aromatic patch of TFE to alanines was performed using a two-step PCR approach. pTFE "Q" mutagenic primers and WT primers. The template for mutagenesis was an overexpression plasmid (pET21b) containing the wild type TFE gene from *Pyrococcus furiosus.* The mutagenic primers were purchased from Integrated DNA technologies and diluted to 100 μmolar (100 pmol/μl) final concentration and placed in a -20°C freezer per manufacturer recommendation. Two initial PCR reactions were performed using a mutagenic primer paired with a WT primer. The mutagenic primers contained 5' ends with mutations changing amino acids 76-WYYYYW-81 to 76- AAAAAA-81. The reaction set up is as follows: 1 ng of TFE plasmid, dNTPs, 1x thermopol buffer and 100 pmol of each oligonucleotide primers. A touchdown PCR was performed using the following parameters: the initial step was at 94°C for 5 minutes followed by a 30 second step at 94°C. The annealing temperature was 56°C for 30 seconds followed by an extension step at 72°C. This was repeated for 4 cycles. The next set of steps was the same with the exception of the annealing temperature, which was decreased to 55°C. These steps were repeated for 4 cycles. The final set of steps decreased the annealing temperature to 54°C. These initial PCR products were gel purified using an 8% PAGE. The samples were removed from the gel using a UV light to identify the bands. The bands were cut out of the gel and eluted

overnight at 37°C in TE' buffer. The eluted DNA was subjected to an ethanol wash, quantified, and cleaned with a Fermentas GeneJet PCR purification kit following manufacturers recommendations.

These PCR products, which have ends overlapping the mutagenic region of the protein, were then used in a second PCR reaction in which 6 individual reactions were set up with 50ng, 25ng, 5ng, and 1ng respectively of each product along with upper and lower WT primers. The resulting products were quantified and purified using the Fermentas GeneJet purification kit. These PCR products were subjected to a restriction digest with NcoI, BamHI, and EcoRI restriction enzymes, along with pET21b-$H_6$-NcoI vector. Following restriction digest the products were ligated into the vector at a 4:1 plasmid to insert ratio. Ligated products were subjected to a phenol extraction/ETOH wash and re-suspended in 5μL of $ddH_2O$. 5μL of this DNA was subjected to a drop dialysis and transformed into *E. coli* electrocompetent cells. Transformants were screened for presence of insert by sequencing. Two colonies demonstrating successful ligation were grown in 10mL of LB broth with 100mg/mL of ampicillin on a 37°C shaker bath overnight. The plasmid was purified from the *E. coli* using the Fermentas GeneJet plasmid miniprep kit. The mutation was confirmed by sequencing and the plasmid containing the mutation was inserted into BL21 gold *E. coli* cells for protein overexpression.

Site directed mutagenesis was performed to mutate the individual aromatic amino acids into alanines using primers purchased from Integrated DNA technologies diluted to100μmolar (100 pmol/μl) final concentration by addition of double distilled water and

placed in a -20°C freezer per manufacturer recommendation.

Primers used for mutagenesis are as follows. The primers are shaded and the wildtype sequence is not. The position of the mutations indicated by the lack of connecting lines.

Figure 16:

```
F66A              5'-acgatgctaagcttgcaaccgctagaagagttagagatgacg-3'
                     ||||||||||||||||||||  ||||||||||||||||||||
                  acatgctacgattcgaacgttggaaatcttctcaatctctactgctct

F66A_antisense    tgtacgatgctaagcttgcaacctttagaagagttagagatgacgaga
                     ||||||||||||||||||||
                  |||||||||||||||||||| 3'-
                  tgctacgattcgaacgttggcgatcttctcaatctctactgc-5'

R70A              5'-cttgcaacctttagaagagttgcagatgacgagactggttg-3'
                     ||||||||||||||||||||  ||||||||||||||||||
                  ttcgaacgttggaaatcttctcaatctctactgctctgaccaaccat

R70A_antisense    aagcttgcaacctttagaagagttagagatgacgagactggttggta
                     ||||||||||||||||||||
                  |||||||||||||||||| 3'-
                  gaacgttggaaatcttctcaacgtctactgctctgaccaac-5'

W76A              5'-gttagagatgacgagactggtgcgtattattattactggcgcat-3'
                     ||||||||||||||||||||  ||||||||||||||||||||
                  tctcaatctctactgctctgaccaaccataataataatgaccgcgtaact

W76A_antisense    agagttagagatgacgagactggttggtattattattactggcgcattga
                     ||||||||||||||||||||
                  |||||||||||||||||||| 3'-
                  caatctctactgctctgaccacgcataataataatgaccgcgta-5'

Y77A              5'-gagatgacgagactggttgggcttattattactggcgcattg-3'
                     ||||||||||||||||||||  ||||||||||||||||||||
                  aatctctactgctctgaccaaccataataataatgaccgcgtaactat

Y77A_antisense    ttagagatgacgagactggttggtattattattactggcgcattgata
                     ||||||||||||||||||||
                  |||||||||||||||||||| 3'-
                  ctctactgctctgaccaacccgaataataatgaccgcgtaac-5'

Y78A              5'-gatgacgagactggttggtatgcttattactggcgcattgatac-3'
                     ||||||||||||||||||||  ||||||||||||||||||||
                  tctctactgctctgaccaaccataataataatgaccgcgtaactatgatt
```


Y78A_antisense
agagatgacgagactggttggtattattattactggcgcattgatactaa
3'-
ctactgctctgaccaaccatacgaataatgaccgcgtaactatg-5'
Y79A
5'-agagatgacgagactggttggtattatgcttactggcgcattga-3'
caatctctactgctctgaccaaccataataataatgaccgcgtaactatg
Y79A_antisense
gttagagatgacgagactggttggtattattattactggcgcattgatac
3'-
tctctactgctctgaccaaccataatacgaatgaccgcgt
aact-5'
Y80A
5'-agatgacgagactggttggtattattatgcctggcgcattgatac-3'
atctctactgctctgaccaaccataataataatgaccgcgtaactatgatt
Y80A_antisense
tagagatgacgagactggttggtattattattactggcgcattgatactaa
3'-tctactgctctgaccaaccataataatacggaccgcgtaactatg-5'
W81A
5'-gacgagactggttggtattattattacgcgcgcattgatactaaaagat-3'
ctactgctctgaccaaccataataataatgaccgcgtaactatgattttctaatg
W81A_antisense
gatgacgagactggttggtattattattactggcgcattgatactaaaagattac
3'-
ctgctctgaccaaccataataataatgcgcgcgtaactatgattt
tcta-5'
WYY A
5'-
acctttagaagagttagagatgacgagactggtgcggctgcttattactggcgcattgatactaaaa
gattacc-3'
cgttggaaatcttctcaatctctactgctctgaccaaccataataataatgaccgcgtaactatgattt
tctaatgg cct
WYYA_antisense
gcaacctttagaagagttagagatgacgagactggttggtattattattactggcgcattgatactaaaaga
ttaccgga
3'-
tggaaatcttctcaatctctactgctctgaccacgccgacgaataatgaccgcgtaactatgattttctaatgg-5’


Figure 16: Primers and sequences used in mutagenesis.

Mutagenic PCR reactions were performed using 50 ng of pTFE ; an overexpression plasmid (pET21b H6-Nco) with the TFE gene inserted as the template. A blue/white screen was performed as a control to determine efficiency of mutagenic plasmid generation using pWhitescript 4.5 plasmid (pWS) from Stratagene. A PCR reaction for each mutation was prepared with 5 µl of 10x Phusion HF buffer 50ng pTFE"Q" 125 ng of each forward and reverse primer, 1.5 µl of dNTP mix, and 125 ng of Phusion DNA polymerase. A thermocycler was set up for an initial step at 98°C for 30 seconds followed by 18 cycles of 98°C denaturing for 30seconds, a 55°C primer-annealing step for 60 seconds, and an extension step at 68°C for 12 minutes. This was followed by a final extension of 15 minutes at 68°C. To remove wild type plasmid DNA, samples were digested with 2µl of DpnI at 37°C for a period of 2 hours. The DNA was cleaned using a phenol/chloroform extraction and ethanol precipitation. The DNA pellets were dissolved in dd$H_20$. Drop dialysis was performed to remove excess salt using 25mL of Tris buffer pH 8 in a petri dish. A Millipore membrane (47millimeter, 0.25µm) was

placed in the petri dish and allowed to equilibrate for 2 minutes. 4μl of each sample was then removed and placed in a 1.5 mL microcentrifuge tube along with 40μl of *E. coli* electrocompetent cells. After mixing this was transferred to a chilled electroporation cuvette. The samples were then transformed using an electroporator set to 200 ″, 25 μF, 51and 1.25 kV for 4.5 seconds. Samples were plated out on LB-ampicillin plates. Transformants were screened by sequencing.

**3.2 Protein overexpression and Purification.**

*E. coli* overexpression strain Bl21 "gold" (Agilent technologies) containing mutant plasmids were streaked out on LB plates supplemented with 100μg/mL ampicillin and grown overnight at 37°C. Three colonies from each plate were inoculated in 60mL each of 2XYT culture medium containing 100μg/mL ampicillin. These were grown at 37°C to an $A_{600}$ optical density of 1.0. When the cell density reached 1.0, IPTG was added to induce overexpression of selected mutants. After 2 hours the cells were placed in chilled Oakridge tubes and spun in an SS34 centrifuge for 5 minutes at 10,000 rpm. The supernatant was decanted and the cells were suspended in lysis buffer containing lysozyme and incubated on ice. Following incubation the samples were then sonicated 3 times at 10 seconds each with a two-minute rest period between bursts.

A his-tag purification protocol was used to isolate the wild type and mutant TFE from other cellular components. The sonicated samples were spun in an SS34 centrifuge for 30 minutes at 14,000 rpm. The supernatant was run through a Ni-NTA column. Each protein containsed a 6xhistidine tag that bound to the column and allows for its separation. Once bound the protein was subjected to several washes and fractions were

collected. The buffers used in washing contained increasing amounts of imidazole, which acts as a competitor for binding to the Ni-NTA matrix.

The initial fractions were collected as flow through (FT). The subsequent fractions were collected: wash 1 (W1), after application of 300µL of a buffer containing10mM imidazole, wash 2 (W2), fraction was collected after 300µL of a buffer with 20mM of imidazole was applied to the column, and elutions 1 and 2 (E1 and E2), were collected after two subsequent applications of 300µL each of a buffer with 200mM of imidazole.

TFB2 was also purified using a Ni-NTA column with the following modifications to the above protocol: W1 and W2 also contained 8M urea, wash 3 (W3) contained 3M urea, wash 4 (W4), E1, and E2 contained 0M urea. Any remaining urea was removed through microcentrifugation and washing with a buffer containing no urea.

### 3.3 *In vitro* transcription assays

*In vitro* transcription assays were performed using WT and mutant TFE proteins. The 5x transcription buffer contains 40mM Na-HEPES pH7.3, 250mM NaCl, 2.5mM $MgCl_2$, 0.1mM EDTA, 5mM beta-mercaptoethanol, 5% (v/v) glycerol, and 0.1µg/µL bovine serum albumin (BSA) in a final reaction volume of 12.5µL. 10nM of gdh promoter DNA was combined with 10nM recombinant RNA polymerase provided by the Thomm laboratory in Germany, 60nM TBP from the Bartlett lab purified through a his- tag purification protocol, 120nM TFB1 or TFB2, and 200nM TFE variant also purified through a his-tag purification protocol. All components were combined,

covered with 30μL of mineral oil and incubated at 65°C for 20 minutes. To ameliorate non-specific interactions between the polymerase and DNA, Heparin was added at a final concentration of 50μg/μL and incubated or 30 seconds. Ribonucleotide triphosphates were added to a final concentration of 500μM ATP, CTP, and GTP, and 10μM UTP orfor low concentration transcription assays, 20μM final of ATP, CTP, and GTP. 10uM of [$\alpha^{-32P}$] UTP [~40Cimmol$^{-1}$]) was also added. Samples were incubated for an additional 20 minutes. 12.5μL of a stop buffer consisting of 8M urea, 0.05M EDTA, 0.09M Tris-borate buffer, 0.02% xylene cyanol and 0.02% bromophenol blue to stop the reaction. The samples were then loaded on a 10% polyacrylamide urea gel. The gel was dried and exposed to a phosphorimager plate. The plate was then scanned on a Typhoon scanner and the bands visualized using Image Quant TL software. Protein was quantified by using Image Quant software by comparing proteins to BSA standards.

# CHAPTER 4

## DISCUSSION

### 4.1 The role of TFE in the transcription cycle

All sequenced archaeal genomes encode a gene for TFE and attempts at deletion result in a lethal phenotype *(Thomm lab personal communication).* Sequence alignments of TFE across archaeal and eukaryotic species demonstrate a conserved patch of aromatic amino acids in the tip of the winged helix (Fig.3).

TFE is not required for transcription *in vitro* although it was shown to partially alleviate defective transcription in the presence of TFB2 *(Micorescu et al., 2008)*. It has also been shown to enhance transcription at some promoters and stimulate transcription at low TBP concentrations *(Bell, Brinkman, van der Oost, & Jackson, 2001).* The focus of the research presented here was to further elucidate the function of TFE in *P. furiosus* and its role in transcription through mutations in and around the highly conserved aromatic patch of the winged helix.

Structural and biochemical experiments have demonstrated that TFE enhances double stranded DNA separation during transcription initiation through an unknown mechanism. TFE also interacts with the RNA polymerase clamp coiled coil through its winged helix *(Grohmann & Werner, 2011)*. In bacteria, sigma 70 region 2.3 contains conserved aromatic amino acids that interact with DNA to assist in promoter opening *(Feklistov & Darst, 2011)*.

Photochemical crosslinking experiments have demonstrated interaction between the W76 and Y78 of the winged helix and the non-transcribed strand at the -9 position

(*Brown and Bartlett, unpublished 2013).* It is likely that TFE interacts with the promoter non-transcribed strand in a similar manner to sigma region 2.3.
It was proposed that due to its interaction with non-transcribed strand transcription would be severely hampered by mutation of this aromatic patch.

In standard transcription assays performed with nucleotide triphosphates in excess using the gdh promoter, TFE with point mutants in the aromatic patch no matter whether TFB1 or TFB2 were used did not show effects on the efficiency of PIC formation. Assays for TFE activity with TFB2 demonstrate that a substitution of the entire aromatic patch shows results in transcription levels similar to that of TFB2 alone and also that the presence of WT TFE enhances transcription minimally (Figure 5). Thus, this assay is not appropriate to determine the role of individual mutations, mainly because under the conditions used here that WT TFE is not required for efficient transcription *in vitro.* It is not clear why TFE has such a small effect on transcription under standard *in vitro* conditions. The simple explanation is that the step it normally affects *in vivo* is not rate limiting under these conditions.

In an effort to establish an assay to elucidate the function of TFE, transcription complex stability was assessed by the addition of excess heparin, a polyanion that can sequester RNA polymerase. The hypothesis tested was that the transcription complexes formed with TFB2 was less stable than those formed with TFB1 and therefore be more susceptible to high concentrations of heparin. The presence of TFE could then provide stability to the PICs formed with TFB2 resulting in higher transcription. The results show that PICs formed with TFB1 and TFB2 respectively and challenged with 1,000 mg/mL of

heparin demonstrate a decrease in transcription compared to PICs formed with TFB1 and TFB2 and the addition of TFE does not increase transcription (see figure 12). Therefore, I conclude that TFE does not stabilize the preinitiation complexes under the conditions tested.

Recent comparisons between the bacteria and eukaryotic transcriptional systems demonstrate similar topology between the TFIIB b-reader and TFB and sigma 70 region 3.2 *(Sainsbury, Niesser, & Cramer, 2013)*. Pupov et al 2014, performed experiments to elucidate the function of sigma region 3.2 in the bacteria *Escherichia coli* by making mutants made in the region of *E. coli* sigma 70 region 3.2. This region can be considered analogous to the divergent sequences of TFB2 in the B-reader. The function of sigma 3.2 was tested in transcription assays performed with high (200μM ATP, CTP, GTP, and 10μM UTP) versus low (20μM ATP, CTP, GTP and 10μM UTP) NTP concentrations. Their results demonstrated that higher NTP concentrations overcome the presence of mutations in sigma 3.2 and were indistinguishable from WT sigma. As demonstrated in figure 14 transcription complexes containing TFB2 are more susceptible to low nucleotide concentrations demonstrating highly deficient transcription. In effect, the divergent sequences of TFB2 cause it to behave similarly to sigma region 3.2 mutants in response to low nucleotide levels. Interestingly, the presence of WT TFE overcame this deficiency (Figure 14). The six alanine substitution in TFE prevents rescue of TFB2 in low nucleotide conditions. Further analysis demonstrates that the presence of TFE increased transcription by 3 fold in assays performed using TFB2 at low nucleotide concentrations. In addition, assays performed

using the six-alanine substitution demonstrated an increase in transcription but only 1.28 fold, which is also comparable to the effects seen under high nucleotide concentrations. This implies an important role for the aromatic patch of the TFE winged helix in transcription initiation, although the mutation seems not to entirely eliminate the functionality of the protein. Furthermore, this analysis reveals an intriguing connection between the efficiency of open complex formation and the response to NTP levels. This implies that under low nucleotide concentrations transcription may be facilitated by TFE with preinitiation complexes formed with TFB2.

### 4.2 The interaction between TFE and TFB2

The positioning of archaeal TFBs in transcription initiation complexes can be inferred by comparison with eukaryotic TFIIB/RNAP II structural models and through cross-linking data.

Crystal structures of RNA polymerase II and TFIIB from *Saccharomyces cerevisiae* with DNA reveal that TFIIB closes the Pol II cleft and positions the DNA inside the cleft to assist in separating the two strands *(Sainsbury et al., 2013).* The B-reader binds the upstream T-strand and aids in recognition of the initiator promoter element and in positioning of the transcription start site *(Sainsbury et al., 2013).* TFIIB repositions active site residues, induces binding of the catalytic magnesium ion, and stimulates formation of the nascent RNA *(Sainsbury et al., 2013).* TFIIB then inhibits the RNA from tilting, which could constrain RNA synthesis *(Sainsbury et al., 2013).* After the addition of 6 nucleotides the nascent RNA is directed toward the exit tunnel and after 12-13 nucleotides there is a steric clash between the RNA and TFIIB, which

causes displacement of TFIIB and formation of the elongation complex *(Sainsbury et al., 2013).*

Recent comparisons between the crystal structures of an RNAP-TFIIB DNA complex from *Saccharomyces cerevisiae* with bacterial RNAP and sigma illustrates a TFIIB like topology and a B-reader-like loop in region 3.2 of sigma. This suggests an evolutionary link between the three homologs *(Sainsbury et al., 2013).*

Further comparisons between TFIIB, TFB, and sigma region 3.2 show conservation of key motifs between the three domains of life *(Taylor et al., 2014).* The implication is that the mechanism for all transcription initiation is evolutionarily conserved *(Sainsbury et al., 2013)*.

Crosslinking experiments from the Bartlett lab demonstrate the position of the C-terminus of TFB2 is oriented similarly to that of TFB1, but the N-terminus is oriented differently and that TFE assists in repositioning of the N-terminus *(Micorescu et al., 2008 and Bhattarai MS Thesis 2014*). Crosslinking results also determined that TFE interacts with the N-terminus of TFB, but not the C-terminus *(Grunberg et al., 2007)*

Bacterial sigma with mutated region 3.2 is analogous to TFB2 in which the B-linker and B-reader are missing key residues, therefore experiments in which a lowered concentration of NTPs demonstrates highly deficient transcription in *P. furiosus.* The experiments here demonstrate that wild type TFE can compensate for low NTP concentration in the presence of TFB2 possibly by stabilizing the open complex thus allowing for a greater opportunity for NTPs to form an initiating complex. This assay provides a much-needed window into studies of TFE function via mutagenesis.

### 4.3 Future experiments

To establish the importance of individual amino acids in the TFE aromatic patch in transcription all of the point mutants, in addition to the two triple mutations should be used in transcription assays under low nucleotide conditions. These experiments will help determine whether one or a few aromatic residues are critical for TFE function, or whether TFE requires the entire intact aromatic patch for function.

As described earlier, there is little need for TFE in transcription *in vitro* using the gdh promoter. One possibility is that open complexes form easily on linear DNA templates so that TFE has nothing to contribute to the transcription mechanism. All genomes of hyperthermophiles contain a gene for reverse gyrase, which is responsible for the positive supercoiling of DNA *in vitro* this is thought to be a mechanism for protection of the genome at high temperatures *(Rodriguez & Stock, 2002)*. However, the positive supercoils could cause an inhibition of promoter opening because of DNA overwinding. It is possible that TFE can assist in changing the topology of DNA in preparation for transcription. To test this promoter DNA templates can be treated with reverse gyrase to generate positive supercoils, and these can be used in assays with WT TFE and TFE aromatic patch mutants.

To answer the question as to whether any of the non-functional mutants can even join the transcription complex, gel shift assays or photochemical cross-linking can be used. Cross-linking can also be used to determine if TFE is repositioning TFB1 and TFB2 in the transcription complex and if the aromatic patch plays a role in this.

TFE and the elongation factor Spt4/5 compete for binding at the clamp coiled

coil during the transition from initiation to elongation *(Grohmann et al., 2011)*. It is possible that the winged helix mutant TFEs may have decreased ability to compete for binding in the presence of Spt4/5. Transcription assays performed with Spt4/5 and mutant TFEs can aid in determining if the aromatic patch plays a role in the transition from initiation to elongation.

Attempts at deletion of TFE have resulted in a lethal phenotype. However, it is possible that subtle mutations within the TFE activation patch could be introduced into cells without causing lethality. Phenotypes associated with such mutants could be tested, for instance, by looking at transcription promoter reporter fusions. If the mutants survive but are sick, suppressor mutations could be isolated which could help identify genetic interactions along with TFE and promoter opening *in vivo.* Therefore, it would be of benefit to introduce the gene for the mutant TFE into a living cell to determine if the mutant has an effect on cell physiology.

My results suggest a role for TFE under low nucleotide conditions in assays with TFB2. It is possible that TFE provides a stabilizing effect that cannot be visualized under high nucleotide conditions and/or there is a repositioning of TFB1 and TFB2 in the presence of TFE that can be further investigated using photochemical cross linking. This assay provides an avenue for further investigation of TFE and the TFE mutants.

## REFERENCES


Aravind, L., & Koonin, E. V. (1999). DNA-binding proteins and evolution of transcription regulation in the archaea. *Nucleic Acids Research*, *27*, 4658–4670. doi:gkc632 [pii]

Arnold, K., Bordoli, L., Kopp, J., & Schwede, T. (2006). The SWISS-MODEL workspace: A web-based environment for protein structure homology modelling. *Bioinformatics*, *22*, 195–201. doi:10.1093/bioinformatics/bti770

Basu, R. S., Warner, B. A., Molodtsov, V., Pupov, D., Esyunina, D., Fernandez-Tornero, C., … Murakami, K. S. (2014). Structural Basis of Transcription Initiation by Bacterial RNA Polymerase holoenzyme. *Journal of Biological Chemistry*. 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. *EMBO Rep*, *2*(2), 133–8. Retrieved fro http://embor.embopress.org/content/embor/2/2/133.full.pdf

Bleiholder, A., Frommherz, R., Teufel, K., & Pfeifer, F. (2012). Expression of multiple tfb genes in different Halobacterium salinarum strains and interaction of TFB with transcriptional activator GvpE. *Arch Microbiol*, *194*(4), 269–279. doi:10.1007/s00203-011-0756-z

Charoensawan, V., Wilson, D., & Teichmann, S. A. (2010). Lineage-specific expansion of DNA-binding transcription factor families. *Trends Genet*, *26*(9), 388–393. doi:10.1016/j.tig.2010.06.004

Debnath, S., Roy, N. S., Bera, I., Ghoshal, N., & Roy, S. (2013). Indirect read-out of the promoter DNA by RNA polymerase in the closed complex. *Nucleic Acids Research*, *41*(1), 366–377. doi:10.1093/nar/gks1018

DeDecker, B. S., O'Brien, R., Fleming, P. J., Geiger, J. H., Jackson, S. P., & Sigler, P. B. (1996). The crystal structure of a hyperthermophilic archaeal TATA-box binding protein. *Journal of Molecular Biology*, *264*(5), 1072–84. doi:10.1006/jmbi.1996.0697

Feklistov, A., & Darst, S. A. (2011). Structural basis for promoter-10 element recognition by the bacterial RNA polymerase sigma subunit. *Cell.*, *147*(6), 1257–1269. doi:10.1016/j.cell.2011.10.041

Gietl, A., Holzmeister, P., Blombach, F., Schulz, S., von Voithenberg, L. V, Lamb, D.C.,
Grohmann, D. (2014). Eukaryotic and archaeal TBP and TFB/TF(II)B follow different promoter DNA bending pathways. *Nucleic Acids Research.* doi:10.1093/nar/gku273

Grohmann, D., Nagy, J., Chakraborty, A., Klose, D., Fielden, D., Ebright, R. H., … Werner, F. (2011). The initiation factor TFE and the elongation factor Spt4/5 compete for the RNAP clamp during transcription initiation and elongation. *Mol Cell*, *43*(2), 263–274. doi:S1097-2765(11)00427-8 [pii] 10.1016/j.molcel.2011.05.030

Grohmann, D., & Werner, F. (2011). Recent advances in the understanding of archaeal transcription. *Curr Opin Microbiol*, *14*(3), 328–334. doi:S1369-5274(11)00060-9 [pii] 10.1016/j.mib.2011.04.012

Grunberg, S., Bartlett, M. S., Naji, S., & Thomm, M. (2007). Transcription factor E is a part of transcription elongation complexes. *The Journal of Biological Chemistry.*, *282*(49), 35482–35490.

Grunberg, S., & Hahn, S. (2013). Structural insights into transcription initiation by RNA polymerase II. *Trends Biochem Sci*, *38*(12), 603–611. doi:10.1016/j.tibs.2013.09.002

Grunberg, S., Warfield, L., & Hahn, S. (2012). Architecture of the RNA polymerase II preinitiation complex and mechanism of ATP-dependent promoter opening. *Nat Struct Mol Biol*, *19*(8), 788–796. doi:10.1038/nsmb.2334

Hahn, S. "Structure and Mechanism of the RNA Polymerase II Transcription Machinery." *Nat Struct Mol Biol* 11.5 (2004): 394–403. Web.

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. *The Journal of Biological Chemistry.*, *271(47)*, 30144–30148.

He, Y., Fang, J., Taatjes, D. J., & Nogales, E. (2013). Structural visualization of key steps in human transcription initiation. *Nature.*, *495*(7442), 481–486. doi:10.1038/nature11991

Hirata, A., Klein, B. J., & Murakami, K. S. (2008). The X-ray crystal structure of RNA polymerase from Archaea. *Nature.*, *451*(7180), 851–854. doi:nature06530 [pii] 10.1038/nature06530

Kiefer, F., Arnold, K., Künzli, M., Bordoli, L., & Schwede, T. (2009). The SWISS-MODEL Repository and associated resources. *Nucleic Acids Research*, *37*. doi:10.1093/nar/gkn750

Korkhin, Y., Unligil, U. M., Littlefield, O., Nelson, P. J., Stuart, D. I., Sigler, P. B., … Abrescia, N. G. (2009). Evolution of complex RNA polymerases: the complete archaeal RNA polymerase structure. *PLoS Biol*, *7*(5), e1000102. doi:10.1371/journal.pbio.1000102

Kostrewa, D., Zeller, M. E., Armache, K. J., Seizl, M., Leike, K., Thomm, M., & Cramer, P. (2009). RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature.*, *462*(7271), 323–330. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt= Citation&list_uids=19820686

Kulbachinskiy, A., & Mustaev, A. (2006). Region 3.2 of the sigma subunit contributes to the binding of the 3'-initiating nucleotide in the RNA polymerase active center and facilitates promoter clearance during initiation. *The Journal of Biological Chemistry.*, *281*(27), 18273–18276. doi:10.1074/jbc.C600060200

Littlefield, O., Korkhin, Y., & Sigler, P. B. (1999). The structural basis for the oriented assembly of a TBP/TFB/promoter complex. *Proceedings of the National Academy of Sciences of the United States of America.*, *96*(24), 13668–13673. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC24122/pdf/pq013668.pdf

Magill, C P, S P Jackson, and S D Bell. "Identification of a Conserved Archaeal RNA Polymerase Subunit Contacted by the Basal Transcription Factor TFB." *The Journal of biological chemistry.* 276.50 (2001): 46693–6. Web.

Meinhart, A., Blobel, J., & Cramer, P. (2003). An extended winged helix domain in general transcription factor E/IIE alpha. *The Journal of Biological Chemistry.*, *278*(48), 48267–48274. doi:10.1074/jbc.M307874200

Micorescu, M., Grunberg, S., Franke, A., Cramer, P., Thomm, M., & Bartlett, M. (2008).
Archaeal transcription: function of an alternative transcription factor B from Pyrococcus furiosus. *Journal of Bacteriology.*, *190*(1), 157–167.

Molodtsov, V., Nawarathne, I. N., Scharf, N. T., Kirchhoff, P. D., Showalter, H. D., Garcia, G. A., & Murakami, K. S. (2013). X-ray crystal structures of the Escherichia coli RNA polymerase in complex with benzoxazinorifamycins. *J Med Chem*, *56*(11), 4758–4763. doi:10.1021/jm4004889

Murakami, K. S., & Darst, S. A. (2003). Bacterial RNA polymerases: the wholo story. *Curr Opin Struct Biol*, *13*(1), 31–39. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt= Citation&list_uids=12581657

Nelson, D. L., & Cox, M. M. (2005). *Lehninger Principles of Biochemistry*. *Biological Membranes* (Vol. 4th editio). doi:10.1016/j.jse.2011.03.016

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. *Molecular and Cellular Biology.*, *15*(9), 4856–4866. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7651404

Ohkuma, Y., & Roeder, R. G. (1994). Regulation of TFIIH ATPase and kinase activities by TFIIE during active initiation complex formation. *Nature.*, *368*(6467), 160–163. doi:10.1038/368160a0

Okuda, M., Tanaka, A., Satoh, M., Mizuta, S., Takazawa, M., Ohkuma, Y., & Nishimura,
Y. (2008). Structural insight into the TFIIE-TFIIH interaction: TFIIE and p53 share the binding region on TFIIH. *The EMBO Journal.*, *27*(7), 1161–1171. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt= Citation&list_uids=18354501

Pupov, D., Kuzin, I., Bass, I., & Kulbachinskiy, A. (2014). Distinct functions of the RNA polymerase sigma subunit region 3.2 in RNA priming and promoter escape. *Nucleic Acids Research.* doi:10.1093/nar/gkt1384

Rodriguez, A. C., & Stock, D. (2002). Crystal structure of reverse gyrase: insights into the positive supercoiling of DNA. *The EMBO Journal.*, *21*(3), 418–426. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11823434

Sainsbury, S., Niesser, J., & Cramer, P. (2013). Structure and function of the initially transcribing RNA polymerase II-TFIIB complex. *Nature.*, *493*(7432), 437–440. doi:10.1038/nature11715

Santangelo, T. J., Cubonova, L., James, C. L., & Reeve, J. N. (2007). TFB1 or TFB2 is sufficient for Thermococcus kodakaraensis viability and for basal transcription in vitro. *Journal of Molecular Biology.*, *367*(2), 344–357. doi:10.1016/j.jmb.2006.12.069

Soppa, J., & Universitèat Frankfurt, B. N. I. fèur M. G. soppa em uni-frankfurt de.

(1999). Transcription initiation in Archaea: facts, factors and future aspects. *Molecular Microbiology.*, *31(5)*, 1295–1305.

Stetter, K. O. (2006). Hyperthermophiles in the history of life. *Philos Trans R Soc Lond B Biol Sci*, *361*(1474), 1833–1837. doi:10.1098/rstb.2006.1907

Tanaka, Aki et al. "Association of the Winged Helix Motif of the TFIIE a Subunit of TFIIE with Either the TFIIE B Subunit or TFIIB Distinguishes Its Functions in Transcription." (2014): 1–14. Web.

Taylor, P., Burton, S. P., Burton, Z. F., (2014). eukaryotic TFIIB are homologs The s enigma : Bacterial s factors , archaeal TFB and eukaryotic TFIIB are homologs, (December), 37–41. doi:10.4161/21541264.2014.967599

Teichmann, M., Dumay-Odelot, H., & Fribourg, S. (2012). Structural and functional aspects of winged-helix domains at the core of transcription initiation complexes. *Transcription*. doi:10.4161/trns.3.1.18917

Thomm, M. (1996). Archaeal transcription factors and their role in transcription initiation. *FEMS Microbiol Rev*, *18*(2-3), 159–171.

Werner, F., & Grohmann, D. (2011). Evolution of multisubunit RNA polymerases in the three domains of life. *Nat Rev Microbiol*, *9*(2), 85–98. doi:nrmicro2507 [pii] 10.1038/nrmicro2507

http://www.funakoshi.co.jp/data/datasheet/PCC/20332.pdf