„Transcription Regulation in the hyperthermophilic crenarchaeon *Thermoproteus tenax* strain Kra1"

Dissertation

zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
– Dr. rer. nat. –

vorgelegt von

Jeannette Marrero Coto

geboren am 27.09.1977, in Havana

Biofilm Centre, Molekulare Enzymtechnologie und Biochemie
Fachbereich Chemie
der
Universität Duisburg-Essen

2010

Die vorliegende Arbeit wurde im Zeitraum von November 2006 bis Oktober 2009 im Arbeitskreis von Prof. Dr. Bettina Siebers in der Fakultät Chemie, Biofilm Centre, Molekulare Enzymtechnologie und Biochemie der Universität Duisburg-Essen durchgeführt.

Tag der Disputation: 2.6.2010

Gutachter: Prof. Dr. Bettina Siebers
Prof. Dr. Ann Ehrenhofer-Murray
Vorsitzender: Prof. Dr. G. Haberhauer

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

The unexpected diversity of molecular and biochemical properties found in Prokaryotes is mainly due to the archaeal members (WHITMAN et al., 1999). Archaea are widespread in diverse habitats from ocean to soil (DE LONG, 1998). Most culturable Archaea are extremophiles, microbes that not only tolerate, but grow optimally in habitats normally considered too harsh for life, for example, they flourish at high temperature, low or high pH, or high salt. Therefore, Archaea are fundamental components of the biogeochemical cycles on earth and some dominate special ecosystems that are of great interest (WHITMAN et al., 1999). Although several extremophilic microorganisms have been known for more than 40 years, the search for extremophiles has intensified in the last decade for two main reasons. First, the range of conditions under which life can exist is now known to be much broader than previously thought, and this has led to the exploration of many hitherto uninvestigated habitats. Second, it is now recognized that attributes of organisms adapted to extreme environments or their enzymes have the potential to serve a wide range of industrial applications. The explosively increasing number of complete genome sequences, especially for hyperthermophiles, enables comparative analyses of the general metabolic capacity and identification of functions of archaea-specific genes and thereby allows to gain a better understanding of cellular processes in Archaea (SCHIRALDI et al., 2002).

To date, it is accepted that Archaea and Bacteria are the two independent phylogenetic lineages of Prokaryotes (WOESE & FOX, 1977). The assignment of Archaea as the third domain of life, in addition to Bacteria and Eukaryotes, based on universal small subunit RNA (SSU RNA) and protein trees has been confirmed by comparative analysis. Archaea are divided into four kingdoms: Euryarchaeota and Crenarchaeota, both of which are represented by many species; Korarchaeota, a phylum including exclusively uncultivated species and the ancient archaeal group Nanoarchaeaota, to date represented by the tiny parasitic cells with the smallest genomes of all known prokaryotes (*Nanoarchaeum equitans*) and, the recently proposed phylum Thaumarchaeota represented by the two members, *Nitrosopomilus maritimus* and *Cenarchaeum symbiosum* (BROCHIER-ARMANET et al., 2008).

Archaea possess genes with obvious counterparts in Bacteria, indicating that the two groups have functional similarities. However, they also harbor genes specific for Eukaryotes (GRAHAM et al., 2000). A peculiar characteristic of Archaea seems to be the mosaic composition of their genomes, which contain both bacterial and eukaryotic components that are otherwise exclusive to these two domains (KEELING et al., 1995). For example, protein sequences that mostly resemble general transcription factors in Eukaryotes have been

discovered in Archaea (KLENK & DOOLITTLE, 1994), while bacterial-like transcriptional regulators such as Leucine-responsive regulatory protein family (Lrp) (KYRPIDES & OUZOUNIS, 1995) and the terminator NusA (OUZOUNIS et al. 1995) have also been found (OUZOUNIS & KYRPIDES, 1996). Despite their similarities, Archaea also show exclusive features that distinguishes them from the other two domains, i.e. the stereochemistry of the glycerol backbone of their membrane phospholipids. All known archaeal membrane phospholipids are isoprenoid ethers built on glycerol-1-phosphate (G1P) contrasting to bacterial and eukaryotic membranes composed basically of fatty acid esters linked to the stereoisomer glycerol-3-phosphate (G3P). Furthermore, Archaea harbor several unique metabolic features, e.g. the archaeal central carbohydrate metabolism (CCM) is characterized by unusual pathways and enzymes, many of which differ from their bacterial or eukaryotic counterparts (SIEBERS & SCHÖNHEIT, 2005) and their regulation as well as energetics is not well understood yet.

Interestingly, archaeal informational processing mechanisms share more similarities with Eukaryotes than Bacteria, for example, transcription, protein co-translational targeting and RNA metabolism. The most striking example is the DNA replication apparatus with DNA helicases, primases, polymerases which are completely unrelated to the respective bacterial counterparts (OLSEN & WOESE, 1996). However, some archaeal genetic characteristics are also remarkably similar to Bacteria, such as, the size and organization of their chromosome, the presence of polycistronic transcription units and the presence of Shine-Dalgarno sequences for the initiation of translation. In summary, Archaea are considered as organisms that use eukaryotic-like proteins in a bacterial-like context (BELL & JACKSON, 2001; GRABOWSKI & KELMAN, 2003).

1.1 TRANSCRIPTION IN EUKARYOTES.

Transcription in Eukaryotes is very complex, including tightly coupled processes as mRNA-processing, chromatin condensation, DNA methylation, transcriptional initiation, alternative splicing of RNA, mRNA stability, translational controls, several forms of post-translational modification, intracellular trafficking, and protein degradation (WRAY et al., 2003). Transcriptional initiation appears to be the primary determinant, or one of the most important determinants, of the overall gene expression profile in Eukaryotes (WRAY et al., 2003).

Four different promoter elements have been identified in eukaryotic promoters (LANGRAGE et al., 1998). There are several sequence motifs—which include the TATA box, initiator (Inr),

TFIIB recognition element (BRE), and downstream core promoter element (DPE)—that are commonly found in core promoters (BUTLER & KADONGA, 2002).

There are at least six "basal" or "general" transcription factors (GTFs) present in Eukaryotes, which are required for RNA polymerase (RNAP) recruitment and preinitiation complex (PIC) formation and can assemble into a transcription PIC in the following order: TFIID, TFIIA, TFIIB, RNA polymerase II-TFIIF complex, TFIIE, and then TFIIH. For the purposes of this work, especially the function of TFIID and TFIIB is important. TFIID is a multisubunit protein that consists of TBP (TATA box-binding protein) and approximately 13 TBP-associated factors (TAFs; BURLEY & ROEDER, 1996; ALBRIGHT & TJIAN, 2000; BERK, 2000; VERRIJZER, 2001). There is considerable evidence that TFIID also binds to the Inr in a sequence-specific manner (KAUFMANN & SMALE, 1994; MARTINEZ et al., 1994; BURKE & KADONAGA, 1996; OELGESCHLÄGER et al., 1996). TFIIB is a single polypeptide that interacts with TBP as well as with the DNA upstream of the TATA box (BRE). TFIID and TFIIB are the first two factors that interact with the core promoter. It appears that these two factors have a critical role in the recognition of core promoter motifs (BUTLET & KADONAGA, 2002) in Eukaryotes.

1.2 TRANSCRIPTION IN ARCHAEA.

Gene expression is regulated at the level of transcription initiation also in Bacteria. However, the mechanism of bacterial transcription initiation is different from that directing expression of eukaryotic genes (THOMM, 1996). About 30 years ago evidences began to accumulate indicating that the basal transcription machinery of Archaea is more closely related to that of Eukaryotes than to the bacterial transcription apparatus (i) the large components of the RNA polymerases (RNAPs) are homologous among all domains but a much higher similarity exists between the archaeal and eukaryal versions than between either of these and the bacterial version (ZILLIG et al., 1979; SCHNABEL et al., 1983), (ii) the identification of transcription factors in the methanogenic Euryarchaeon *Methanococcus* (FREY et al., 1990) and the Crenarchaeon *Sulfolobus* (HÜDEPOHL et al., 1990), (iii) the demonstration that archaeal promoters are constituted of a TATA box located 25 bp upstream of the transcription start site, the upstream BRE region and a second essential element located around the start of RNA synthesis named initiator element (Inr) (HAUSNER et al., 1991). This promoter structure resembles the core elements of typical TATA box containing eukaryotic RNA polymerase II (RNAP II) promoters (BREATHNACH & CHAMBON, 1981), (iv) transcription factor encoding genes have been identified in the genomes of Thermococcales and *Sulfolobus* which show

high sequence similarity to the eukaryotic basal transcription factors TATA-binding protein (TBP) and TFIIB (ROWLANDS et al., 1994; WETTACH et al., 1995), (v) eukaryotic TBP from yeast and human can replace an archaeal transcription factor in a *Methanococcus* derived cell-free transcription system (WETTACH et al., 1995); and finally (vi), transcription initiation in Archaea was shown to be mediated by GTFs (GOHL et al., 1995). It seems that the ancestor of Archaea and Eukaryotes evolved from a common basal transcriptional machinery before Archaea diverged from the eukaryotic/archaeal lineage (THOMM, 1996). All this findings motivated the increasing interest in the function of individual components of the archaeal transcriptional machinery and the relationship of archaeal transcription factors to their eukaryotic counterparts.

The minimal archaeal transcriptional machinery consists of a multi-subunit RNAP resembling the eukaryotic RNAP II in sequence and subunit composition (HIRATA et al., 2008; KESSLER et al., 2006; KUSSER et al., 2008; FIORE et al., 2009).), a TATA-box binding protein (TBP) also homologous to its eukaryotic counterpart and transcription factor B (TFB), which is homologous to eukaryotic TFIIB (LANGER et al., 1995). In particular, TBP recognizes and binds to the archaeal promoter comprising of an AT-rich TATA-like element, positioned immediately downstream of the binding site for TFB (i.e., the B recognition element (BRE)). Subsequently, the TBP/TFB/DNA complex recruits RNAP to the promoter to specifically initiate transcription at an initiator sequence (Inr) that is located about 25 bp downstream of the TATA element (BARTLETT, 2005; GEIDUSCHEK & OUHAMMOUCH, 2005). Both TBP and TFB consist of two imperfect direct repeats. In addition, TFB has an N-terminal domain (NTD) forming a Zn-ribbon and a B-finger (GEIDUSCHEK & OUHAMMOUCH, 2005; WERNER & WEINZIERL, 2005; GOEDE et al., 2006).

A third archaeal transcription factor, TFE, is homologous to the N-terminal part of subunit α of eukaryotic TFIIE (BELL et al., 2001; HANZELKA et al., 2001). TFE is not required for promoter directed transcription but can stimulate the activity of some promoters by a factor of three to four. TFE can also complement some mutants of TFB indicating that these proteins interact synergistically and contribute to catalytic core functions of RNAP (GEIDUSCHEK & OUHAMMOUCH, 2005; GOEDE et al., 2006).

An interesting finding is that multiple transcription factor homologues have been identified in several archaeal genomes. It is suggested that different combinations of TFB-TBP might function similar to bacterial sigma factors as part of a mechanism for regulation of gene expression (BALIGA et al., 2000). In general, only a small number of TBP encoding genes has been identified in available archaeal genomes, whereas, a larger number of TFB encoding

genes is present, suggesting a more complex function of TFB in transcription regulation. It has been suggested previously that different classes of promoters use the same TBP but distinct TFB proteins in promoter recognition (BALIGA et al., 2000).

The functional studies of multiple TBP and TFB factor in the model haloarchaeon, *Halobacterium* sp. NRC-1 represent the first in-depth analysis, which confirms that specific TBP-TFB pairs may be required for gene-specific transcription in Archaea (COKER & DASSARMA, 2007). The study reveals that nearly half of the *tbp* and *tfb* genes are non-essential and can be knocked-out without deleterious effects on the cells grown under standard laboratory conditions. Comparison of two of these gene deletion strains, *tbpD*Δ and *tfbA*Δ, revealed that TBP D regulates 15%, TFB A regulates 18%, and TBP D and TFB A together regulate over 10% of the genes in the genome. These results strongly support the prediction of a novel mechanism of gene regulation, where specific TBP-TFB pairs are used for transcription of specific subsets of genes. The finding that two key heat shock genes, *hsp1* and *cctA*, are under transcriptional control of TBP D and TFB A factors, and that both the *tbpD*Δ and *tfbA*Δ mutants are sensitive to elevated temperatures, suggest that these factors regulate expression of genes important for survival at increased temperature in this Haloarchaeon (COKER & DASSARMA, 2007).

The genome of the hyperthermophilic Euryarchaeon *Thermococcus kodakarensis* encodes two TFBs, and it was shown previously that either of which may be deleted without affecting cell growth under laboratory conditions. Both *Kod*-TFBs function in transcription initiation *in vitro*, and no apparent promoter selectivity was observed (SANTANGELO et al., 2007). Also two TFB paralogs, *Pfu*-TFB1 and *Pfu*-TFB2, are encoded by the genome of the closely related member of the Thermococcales *Pyrococcus furiosus*. The *Pfu*-TFB2 N-terminus is not as well conserved; the putative Zn-ribbon-containing portion of the N terminus (amino acids 17 to 49) displays just 45% identity to the *Pfu*-TFB1 Zn-ribbon region (amino acids 7 to 39), and there is no recognizable B-finger motif. *Pfu*-TFB2 transcript levels rise following heat shock, suggesting that *Pfu*-TFB2 polypeptide is involved in the response to heat stress (MICORESCU et al., 2008). Interestingly, *Pfu*-TFB2 functions poorly in promoter-dependent transcription initiation, but photochemical cross-linking experiments indicated that the orientation and occupancy of transcription complexes formed with *Pfu*-TFB2 at the strong glutamate dehydrogenase (*gdh*) gene promoter are similar to the orientation and occupancy of transcription complexes formed with *Pfu*-TFB1. However, initiation complexes formed by *Pfu*-TFB2 display a promoter opening defect that can be bypassed with a preformed

transcription bubble, explaining the observed low *Pfu*-TFB2 transcription activity (MICORESCU et al., 2008).

Generally, crenarchaeal genomes encode one TBP and multiple homologues of TFB. The function of TFB3, the third TFB from *Sulfolobus solfataricus* and one of the most highly upregulated genes following UV irradiation in *S. solfataricus* has been recently proposed (PAYTUBI & WHITE, 2009). *Sso*-TFB3 possesses only half size of TFB1 and was shown to compete with *Sso*-TFB1 for binding to RNAP *in vitro* (FRÖLS et al., 2007) and *Sso*-TFB3 interacts with the RpoK subunit of RNAP and with the ternary complex TFB1/TBP/DNA (PAYTUBI & WHITE, 2009). Moreover, *Sso*-TFB3 stimulates transcription from a variety of promoters *in vitro* in the presence of *Sso*-TBP and *Sso*-TFB1. The dramatic increase in *tfb3* transcript abundance following DNA damage may thus provide a mechanism to modulate general transcription following DNA damage (PAYTUBI & WHITE, 2009). These findings suggest an unusual complexity of transcription regulation in Crenarchaeota, which may in some respect, resemble higher organisms (COKER & DASSARMA, 2007; FAIGER et al., 2006). Interestingly, the presence of TBP-like (DANTONEL et al., 1999) and TBP-related (MALDONADO, 1999) proteins in some metazoan including mammalian cells, respectively, suggests that alternative TBP±TFB±RNAP complexes may be present in eukaryotic systems as well (BURATOWSKI, 1997, BALIGA et al., 2000).

In the group of Crenarchaeota, only TFB1 from *S. solfataricus* and *S. acidocaldarius* (BELL & JACKSON, 2000) and more recently, TFB3 from *S. solfataricus* (PAYTUBI & WHITE, 2009) have been studied biochemically. However, the role of TFB2 in the phylum Crenarchaeota is still unknown and also, no *tfb* knockout mutants are available for any of the crenarchaeal TFBs yet.

1.3 MULTIPROTEIN BRIDGING FACTOR 1 (MBF1).

Relatively little is known about the interaction of transcription regulators with the basal transcription apparatus in Archaea. The majority of archaeal DNA binding proteins with known gene-regulatory function are homologous to bacterial activators and repressors, with only a few transcriptional regulators resembling the eukaryotic ones (ARAVIND & KOONIN, 1999; BELL & JACKSON, 2001). How these “bacterial-like” regulators interact with the “eukaryotic-like” basal transcription machinery of Archaea is still a matter of debate (FIORE et al., 2009).

Most of the studied archaeal regulators are negative regulators and only few activators have been identified. MDR-1 is a transcriptional repressor found in *Archaeoglobus fulgidus* (BELL

et al., 1999) and is homologous to the bacterial metal-dependent transcription repressor DtxR (BOYD et al., 1990). It represses the transcription of the operon comprising its own gene, in addition to genes belonging to an ABC transporter system (BELL et al., 1999). The binding of this repressor does not interfere with the association of TBP or TFB with DNA, but rather interferes with the recruitment of RNAP, which is a mechanism that is characteristic of some bacterial transcriptional repressors (HICKEY et al., 2002).

Archaea also possess homologues of the leucine-responsive regulatory protein (Lrp) (NAPOLI et al., 1999; OUHAMMOUCH & GEIDUSCHECK, 2001), which perform a number of both positive and negative gene-regulatory functions in Bacteria. Lrp influences DNA topology and the formation of DNA-protein complexes in Bacteria (WANG & CALVO, 1993). Lrs14 is an archaeal homologue of Lrp from *S. solfataricus*, and Prt1 and Ptr2 are two Lrp homologues found in *Methanocaldococcus jannaschii* (NAPOLI et al., 1999; OUHAMMOUCH & GEIDUSCHECK, 2001).

Unlike the MDR-1 repressor, Lrs14 and Ptr1 bind DNA and cover the TATA box and the BRE, obstructing the binding of TBP and TFB to the promoter and repressing its own transcription (BELL & JACKSON, 2000). The sequences recognized by Ptr1 and Ptr2 are palindromic, indicating that both molecules interact with DNA as dimers (OUHAMMOUCH & GEIDUSCHECK, 2001; HICKEY et al., 2002). LrpA from *Pyrococcus woesei* inhibits transcription by interfering with RNAP recruitment (DAHLKE & THOMM, 2002; HICKEY et al., 2002).

Mja-Ptr2 serves as positive regulator of transcription: the transcriptional activation is generated from promoter-proximal upstream activating DNA sites (UAS) by recruiting TBP to the promoter (OUHAMMOUCH & GEIDUSCHECK, 2005). So far, only *Mja*-Ptr2 has been shown to activate transcription by its conjugate core transcription apparatus *in vitro*. MDR-1, Lrs14, Prt1 and Prt2 harbor a conserved Helix Turn Helix (HTH) motif (BELL et al., 1999; NAPOLI et al., 1999; OUHAMMOUCH & GEIDUSCHECK, 2001), which is also present in the basal transcription factors TFB and TFE (ARAVIND & KOONIN, 1999), and several prokaryotic regulators (HUFFMAN & BRENNAN, 2002).

Interestingly, multiprotein bridging factor 1 (MBF1) is a transcriptional co-activator, which is present in almost all organisms which harbor TBP as general transcription factor, Archaea and Eukaryotes (ARAVIND & KOONIN, 1999). MBF1 has been identified in all archaeal genomes raising the question about its possible function.

It was previously reported that the genome context of *mbf1* is well conserved in Crenarchaeaota and harbors genes, which encode proteins with predicted functions involved

in informational processing mechanisms such as *pan* (proteasome-activating nucleotidase), *hflX* (G-protein of the HflX family), *tgt* (tRNAguanine transglycosylase), *tfb* and *tfe* (coding for archaeal homologues of transcription factor II B and N-terminal half of alpha – subunit of transcription factor II E, respectively) and *rpoG* (hypothetical RNA-polymerase subunit G) (KONING et al., 2009). In Euryarchaeota, the genomic context of *mbf1* is less conserved; however, proximity to pan is also observed (KONING et al., 2009).

MBF1 was first purified from posterior silk gland extracts of the silkworm *Bombyx mori* and was subsequently found in organisms as distant as mammals, *Arabidopsis* and yeast (TAKEMARU et al., 1997; 1998; DRAGONI et al., 1998; SMITH et al., 1998; KABE et al., 1999; MARIOTTI et al., 2000; GODOY et al., 2001; MATSUSHITA et al., 2002; ZANETTI et al., 2003; JINDRA et al., 2004; TSUDA et al., 2004). From *in vitro* transcription studies it has been suggested that *Bm*-FTZ-F1, a silkworm counterpart of the FTZ-F1 activator in *Drosophila* (LI et al., 1994), activates transcription of the *ftz* gene mediated by MBF1 by binding to the FTZ-F1 site.

In general, MBF1 mediates transcriptional activation by bridging between TBP and either steroid/nuclear hormone receptors (e.g. FTZ-F1 in insects (TAKEMARU et al., 1997, JINDRA et al., 2004), Ad4BP/SF-1 in human (KABE et al., 1999)) or leucine zipper (bZIP)-type transcriptional activators such as GCN4 in yeast (TAKEMARU et al., 1997) and ATF1, c-Jun, as well as c-Fos in human (KABE et al., 1999; KONING et al., 2009). Therefore MBF1-dependent activators are responsible for the regulation of numerous different processes in these organisms (LI et al., 1994; TAKEMARU et al., 1997; 1998; DRAGONI et al., 1998, for recent review see KONING et al., 2009). A unique structural feature of MBF1 interacting activators is that they contain a conserved basic region in their DNA-binding domains. It has previously been suggested that the co-activator MBF1 mediates transcriptional activation by interaction with the conserved basic region, which stimulates the bindings of the co-activator-activator complex to the target DNA sequences, in a similar manner as the human T-cell leukemia virus transactivator Tax does (BARANGER et al., 1995; PERINI et al., 1995; TAKEMARU et al., 1997).

Human MBF1 (hMBF1) binds to both the mammalian ortholog of FTZ-F1, called SF1 or Ad4BP as well as TBP and stimulates transcription *in vivo* in an Ad4BP-dependent fashion (KABE et al., 1999), suggesting that hMBF1 acts as a co-activator similar to its insect (TAKEMARU et al., 1997; LI et al., 1994) and yeast counterparts (TAKEMARU et al., 1998). Two isoforms, which are termed hMBF1α and hMBF1β, have been identified. They share an identical N-terminal region, but the C-terminal part is different (KABE et al., 1999). It has

been predicted that Ad4BP/SF-1-hMBF1 complex plays an important role in the establishment of reproduction in addition to its role in the production of steroid hormones (KABE et al., 1999). In endothelial cells, hMBF1, also known as endothelial differentiation-related factor 1 (EDF1), may have a role in proliferation as antisense inhibition of MBF1 and induces differentiation (DRAGONI et al., 1998), but its function in transcription remains obscure. In addition, hMBF1 is a co-activator for the nuclear receptors LRH1, LXRα, and PPARγ that regulate lipid metabolism (BRENDEL et al., 2002). Apart from its role in transcription, MBF1 is highly expressed in the heart, binds to calmodulin and is phosphorylated by protein kinase C (PKC), which mediates hypertrophic signaling (SMITH et al., 1998; DRAGONI et al., 1998; KABE et al., 1999; MARIOTTI et al., 2000). hMBF1 is also involved in the protection from oxidative stress through the interaction with AP-1. It has been shown that MBF1 protects the critical cysteine residues from oxidation by the direct interaction with D-Jun and stimulates AP-1 binding to DNA (JINDRA et al., 2004). Deletion of *mbf1* causes sensitivity to oxidative stress *in vivo* and compromises an AP-1-dependent process of epithelial tissue closure (JINDRA et al., 2004; MIOTTO & STRUHL, 2006).

The function of MBF1 is also well studied in plants and it has been suggested that plant MBF1 homologues are involved in the defense responses to pathogens. Transcription of potato MBF1 (StMBF1) is up-regulated upon wounding (GODOY et al., 2001) and during fungal attack (ZANETTI et al., 2003). Tobacco MBF1 (NtMBF1a) and two of the three homologues of *Arabidopsis* MBF1s (AtMBF1a and AtMBF1b) have been shown to interact with the tomato mosaic virus movement protein (MATSUSHITA et al., 2002). The expression of *Arabidopsis* MBF1c is elevated in response to pathogen infection, salinity, drought, heat, hydrogen peroxide, and application of the plant hormones abscisic acid or salicylic acid and is a key regulator of thermotolerance (TSUDA et al., 2004; SUZUKI et al., 2008). Moreover, the ER24 gene, a tomato counterpart of MBF1 (LeMBF1), is immediately and transiently induced in ethylene-treated late immature fruit (ZEGZOUTI et al., 1999).

In addition, the role of MBF1 as transcriptional co-activator has been well studied in yeast. yMBF1 is essential for the transcriptional activation of the gene *HIS3*, which encodes the third enzyme of the histidine biosynthesis pathway, imidazole-3-phosphate dehydratase. yMBF1 interacts directly with both the sequence-specific activator (GCN4) and TBP, implying that it also interconnects the bZIP factor with the basal transcription machinery (TAKEMARU et al., 1998).

Apart from the well documented role of MBF1 as transcriptional co-activator in Eukaryotes, the work from HENDRICK and co-workers (2001) suggests that MBF1 could also be involved

in translation fidelity in the yeast mutant *suf13-1*. It was shown that the mutant *suf13-1*, which harbors a modified version of *mbf1* gene, causes suppression of +1 frameshift mutations in the *HIS4*, *LEU2* and *MET2* genes. The predicted peptide produced from *suf13-1* is 80 amino acids long with the first 69 residues derived from wild-type MBF1 sequence followed by 11 residues derived from the sequence corresponding to the -2 reading frame. The authors proposed that the interaction between the truncated product of MBF1 and TBP is impaired affecting the RNA-polymerase III (RNAP III) dependent transcription of tRNA genes. The reduced levels of tRNAs then could induce increased rates of frameshift misreading leading to suppression (HENDRICK et al., 2001). However, up to now, no experimental data supports the connection of MBF1 with RNAP III transcription initiation or translation.

MBF1 contains a Cro-Helix Turn Helix (HTH)-type domain, which is the only highly conserved, classical HTH domain that is vertically inherited in all Archaea and Eukaryotes (ARAVIND & KOONIN, 1999). As described above recent studies revealed diverse biological functions of MBF1 in insects, yeast, rats, plants and human (LI et al., 1994; TAKEMARU et al., 1998; SMITH et al., 1998; KABE et al., 1999; TSUDA et al., 2004), but its function in Archaea remains unclear.

This research work is focussed on transcription regulation in the hyperthermophilic crenarchaeon, *Thermoproteus tenax* strain Kra1, which is a sulfur-dependent anaerobe with optimal growth at 86°C and pH 5 (ZILLIG et al., 1981). The organism is able to grow chemolithoautotrophically on carbon dioxide (CO_2) and hydrogen (H_2) as well as chemoorganoheterotrophically in the presence of different carbon sources (e.g. glucose, starch). The aims of this study are: (i) to unravel the role of multiple general transcription factors (GTFs) in *T. tenax* and to study their function in transcription initiation as well as regulation of basal transcription; (ii) to analyze the role of multiprotein bridging factor 1 (MBF1) as co-activator in Archaea. The present work will not only provide new insights in the regulation of archaeal transcription, but furthermore, will, as a simpler related model, enhance the understanding of transcription and its regulation in Eukaryotes.

2. MATERIALS AND METHODS

2.1 CHEMICALS AND PLASMIDS.

All chemicals and enzymes were purchased from Amersham Pharmacia Biotech Europe GmbH (Uppsala, SWE), Bio-Rad Laboratories GMbH (Munich, GER), Biometra (Göttingen, GER), Difco Laboratories (Augsburg, GER), MEB Fermentas Life Science (St. Leon Rot, GER), Gerbu Biotechnik GmbH (Gaiberg, GER), Life Technologies, (Karlsruhe, GER), Merck, QIAGEN (Hilden, GER), Roche Diagnostics GmbH (Mannheim, GER), Roth GmbH (Karlsruhe, GER), SERVA Electrophoresis GmbH (Heidelberg, GER), Sigma-Aldrich (Taufkirchen, GER), Tropix and VWR International in analytical grade. ^{32}P-γATP was obtained from Hartmann. For heterologous expression the pET vector system (pET-11c, pET-15b, pET-24a (Novagen); pQE-30 (QIAGEN); pET-302 was used.

2.2 INSTRUMENTS.

INSTRUMENT	DESCRIPTION
GelDoc gel documentation	GelDoc Gel Documentation System (Bio-Rad Laboratories GmbH (München, GER) Universal Hood II, Mitsubishi P93DW Digital Monochrome Printer, Malaysie
Fast Flow Liquid Chromatography (FPLC)	Biologic DuoFlow Pathfinder 20 system (Bio-Rad Laboratories GmbH (München, GER) System: F10 work station, MX-1 mixer, 3-Tray rack, AVR7-3 sample inject valve, QuadTec UV/Vis detector with 3 mm PEEK flow cell, system cable 25 (RS-232), BioFrac fraction collector
Chromatography columns	HiLoad 26/60 Superdex 200 prep grade, GE Healthcare (Freiburg, GER), NiNTA prepack column, Bio-Rad Laboratories GmbH (München, GER), Q-sepharose prepacked columns Q-1 and Q-12, Bio-Rad Laboratories GmbH (München, GER)
Autoclaves	H+P Varioklav, 25T, Federgari Autoklave, Integra Bioscience (IBS) (Fernwald, GER) H+P Varioklav, 75S, Federgari Autoklave, Integra Bioscience (IBS) (Fernwald, GER)
Agarose gel electrophoresis system	Agagel Mini, Biometra (Göttingen, GER) Power supply: Consort E143 (MS Laborgeräte)
Easy castelectrophoresis System	B1A/B2 Owl Scientific (Woburn, USA) Power Supply: EPS 30, GE Healthcare (Freiburg, GER)
Vertical gel electrophoresis system	Mini-gel Twin, Biometra GmbH (Göttingen, GER) Power supply: Consort E143 (MS Laborgeräte)
Fermenter	Infors HT Minifors (Bottmingen, Switzerland)
Sonicator	Ultraschallprozessor UP 200s, Oehmen Labortechnik (Hielscher, GER)
French Press	French Pressure Cell, FA-078AE, Thermo Electron Corporation (Milford, USA)

Incubators	Infors HT Unitron (Bottmingen, Switzerland) Heraus B6 Kendro (Langenselbold, GER) Heraus T20 Kendro (Langenselbold, GER)
Photometer	Eppendorf Biophotometer plus, Eppendorf (Hamburg, GER)
Aqua bidest water system	GFL 2104 (Burgwedel, GER)
Thermocycler	Mastercycler personal, Eppendorf (Hamburg, GER) Thermocycler C1000, Bio-Rad Laboratories GmbH (München, GER)
Centrifuges	Sorvall Centrifuge RC26, Kendro (Langenselbold, GER) Rotor: Sorvall SS-34, Kendro (Langenselbold, GER) Rotor: F10S-6x500y, Fiberlite Piramoon Technologies (Santa Clara, USA)
Contamination monitor	Contamat FHT 111M, Thermo Scientific (Erlangen, GER) LB 124 Berthold Technologies GmbH (BadWildbad, GER)
Ultraviolet light	TCP-20.C, Vilbert Lourmat (Marne La Valle, France)
Heat Block	Thermoblock v4.6, Hardware & Service (Frieldland, GER)
Water bath	Thermomix UB, B. Braun (Melsungen, GER)
Sterile work bench	Hera Safe KS12, 1,200x780x627 mm (Langenselbold, GER)
Filtration system	KNF Lab, Laboport (Freiburg, GER)
VersaDoc gel documentation	Versadoc Imaging System, 4000MP, Bio-Rad Laboratories GmbH (München, GER) Tamron AF Aspherical 28 – 80 mm
Microwave	HP1612, Siemens (München, GER)

2.3 STRAINS OF *Escherichia* coli AND GROWTH CONDITIONS.

In this study *Escherichia coli* K-12 DH5α strain (DSMZ 6897) (HANAHAN, 1983), *E. coli* BL21(DE3) (Novagen) (STUDIER & MOFFATT, 1986), *E. coli* BL21-CodonPlus(DE3)-RIL (Stratagene) (CARSTENS & WAESCHE, 1999) and *E. coli* M-15 (Qiagen) were used.

All *E. coli* strains were grown aerobically in 2 – 400 ml batch cultures in reaction tubes or Erlenmeyer flasks at 37°C with aeration by gyratory shaking (180 rpm). Mass cultures of 3.5 L volume were grown in a fermenter aerated by gassing with compressed air through a bacterial tight filter (50 L/min flow rate). Cultures were grown on Lysogeny Broth (LB) medium (Roth). For solid medium plates, 1.5% (w/v) agar-agar (Difco Laboratories, Augsburg, GER) were added. Antibiotics (Sigma-Aldrich, Taufkirchen, GER) were added according to the encoded plasmid resistance in the following concentrations: ampicillin 100 µg/ml, kanamicin 50 µg/ml, chloramphenicol 34 µg/ml. Growth was monitored spectrophotometrically at 600 nm.

E. coli K-12 DH5α was used for the cloning and storage of the derived recombinant vectors. The *E. coli* strains BL21(DE3), BL21-CodonPlus(DE3)-RIL, Rosetta(DE3) and M15 were used for heterologous expression of recombinant *T. tenax* proteins.

2.4 MOLECULAR BIOLOGICAL METHODS WITH DNA.

2.4.1 Preparation of genomic DNA from *T. tenax*.

The isolation of genomic DNA from *T. tenax* was performed using DNAzol reagent (Invitrogen) according to the manufacter´s instructions, with slight modifications. The method is based on the use of a detergent guanidine lysing solution that hydrolyzes RNA and promotes the selective precipitation of DNA from the cell lysate (CHOMCZYNSKI et al., 1993; MACKEY et al., 1996). Briefly, *T. tenax* cells (0.3 g wet weight) were suspended in 2 ml DNAzol reagent and incubated for 15 min at room temperature (RT). The sample was homogenized by using a hand held glass-teflon homogenizer and then incubated for 5 – 10 min at RT. The homogenate was sedimented by centrifugation (10,000 x g, 10 min, RT). The centrifugation step was repeated until all visible traces of sulfur compounds were removed and the viscous supernatant was then transferred to a fresh tube. For DNA precipitation 0.5 ml of ice-cold ethanol 100% (v/v) per 1 ml of DNAzol was added and the sample was mixed by inverting the tube 5 – 8 times and finally incubated at RT for 1 – 3 min. DNA was sedimented by centrifugation (10,000 x g, 10 min, RT). The supernatant was decanted and DNA was washed twice with 1.0 ml ice-cold ethanol 70% (v/v) and then centrifuged again. The remaining ethanol was completely removed via vacuum centrifugation (speed vac) and the pelleted DNA was dissolved in 200 µl A. bidest for 30 min at RT. The quality of purified DNA was checked by restriction digestion and subsequent agarose gel electrophoresis.

2.4.2 Isolation of plasmid DNA from *E. coli*.

Plasmid DNA used for quantitative and qualitative analysis was prepared by the Plasmid Mini kit (MEB Fermentas, St. Leon Rot, GER) following the manufacturer`s instructions. Boiling PCR was used for a rapid qualitative analysis of recombinant *E. coli* clones by PCR-amplification. The colonies were picked with a sterile pipette tip, part of the cells streaked on an LB agar plate and the rest was resuspended in 50 µl A. bidest. Cells were lysed by incubation of the resuspended cells at 94°C for 5 min, centrifuged (14,000 x g, 1 min, RT) and finally 1 µl of the supernatant was used as template DNA for PCR (25 µl volume).

2.4.3 Quantification of DNA.

DNA concentration was measured by determination of absorbance at 260 nm using the Eppendorf spectrophotometer.
The concentration of extracted DNA was determined photometrically at λ = 280 nm. Absorption (A_{280}) = 1, corresponds to 50 µg DNA/ml (SAMBROOK et al., 1989).

2.4.4 Agarose gel electrophoresis.

Agarose gel electrophoresis (SAMBROOK et al., 1989) was employed to monitor the result of a restriction enzyme digestion, to determine the yield or purity of DNA preparations or PCR reactions, and to size fractionate DNA molecules (see below, Purification of DNA fragments). For agarose gels (1%), 1 g (w/v) agarose in TAE-buffer (40 mM Tris-Acetate, 1 mM EDTA) was used.

DNA samples were mixed with loading buffer (6 x concentrated: 0.2% Bromophenol-blue, 0.2% xylencyanol FF, 60% glycerol and 60 mM EDTA) and applied into the sample wells. Electrophoresis was performed at 100V (depending on the distance of electrodes) for 0.5 – 1 hour (RT) for small agarose gels Size marker (GeneRulerTM 1kb DNA ladder, MEB Fermentas, St. Leon Rot, GER) was co-electrophoresed with DNA samples. After electrophoresis the agarose gel was stained in a water bath with 0.2 mg/ml ethidium bromide (Sigma-Aldrich, Taufkirchen, GER) for 20 minutes and destained (20 minutes in a water bath) to enable detection of DNA or DNA fragments under UV light. Documentation was performed by using the GelDoc – Gel Documentation System (BioRad, München, GER).

2.4.5 Purification of DNA Fragments.

Extraction and purification of DNA products and PCR fragments from agarose gels were achieved by using the PCR purification and Gel extraction Kit from Promega (Mannheim, GER) following the manufacter`s instructions.

2.4.6 Amplification of genomic DNA and plasmid DNA by PCR.

For PCR amplification, 50 – 100 ng genomic or plasmid DNA template, 1 µM of each forward and reverse primer, 200 µM dNTPs (Peqlab), 1 µl Go-Taq reaction buffer (Promega, Mannheim, GER) and 1 U of DNA polymerase (Go-Taq polymerase, Promega, Mannheim, GER) were used. Oligonucleotide primers were purchased from Invitrogen. The name and sequence of the primers used in this study are shown in Table 1.

PCRs were performed using a thermocycler. The cycling protocol consisted of 25 - 30 cycles comprising a denaturation step at 95°C for 30 sec, a step for primer annealing at the melting temperature (Tm) for the respective primer set for 30 sec and a step for extension for 1 min per 1,000 bp at 72°C. The melting temperature (Tm) for each primer was calculated using the program Oligo structure ver. 3.4. The final annealing temperature used in the PCR program was 5°C below the calculated temperature for the respective primer set. In all PCR reactions,

TABLE 1. List of primer sets, bacterial expression vectors and expression hosts of *Ttx*-TFBs, *Ttx*-TBP, *Ttx*-MBF1 and yeast MBF1 in *E. coli*.

Gene Name (ORF ID)	Protein	Primer name	Sequence (5'-3'), Restriction site	Expression vector	Host
ttx_tfb1 (ttx_1484)	*Ttx*-TFB1 [c]	TFB1-NdeI-Fwr TFB1-BamHI-Rev	ttataacatatgtcggctca, *Nde*I [b] ttgaggatcctatctgctgg, *Bam*HI [b]	pET11c	Rosetta(DE3)
ttx_tfb2 (ttx_2085)	*Ttx*-TFB2 [c]	TFB2 for6-Fwr TFB2 rev7-Rev	agaggtaggagcatatgactaggcg, *Nde*I [b] aattttgagccaatggatccg, *Bam*HI [b]	pET11c	Rosetta(DE3)
	6xHis-*Ttx*-TFB2	TFB2-Fwr TFB2-Rev	gggggaattctcatgactaggcgaatactagc, *Eco*RI [a] *Bsp*HI [b] ggcgggatcctcatatcgagatctctatatagag, *Bam*HI [a,b]	pET302	Rosetta(DE3)
ttx_tfb3 (ttx_1929)	6xHis-*Ttx*-TFB3	1929-Fwr 1929-Rev	aaagaattccatatgaggtgttcttactgcggctccg, *Eco*RI [a] *Nde*I [b] aaaggatccttagtaattcagctccttgcataatttcc, *Bam*HI [a,b]	pET15b	Rosetta(DE3)
ttx_tbp (ttx_0178)	*Ttx*-TBP [c]	TBP-NdeI-Fwr TBP-BamHI-Rev	aatctacatatggactcttctagagccgcc, *Nde*I [b] ggggacaagtagggggggatccg, *Bam*HI [b]	pET11c	Rosetta(DE3)
ttx_tfb4 (ttx_1732)	6xHis-*Ttx*-TFB4	Tfb4-1732F1 Tfb4-1732R	gggggaattcatatgagtgtacagatagatatag, *Eco*RI [a] *Nde*I [b] ggcgggatcctcacttcacagccctagc, *Bam*HI [a,b]	pET15b	BL21(DE3) Rosetta(DE3)
		Tfb4-1732F2 Tfb4-1732R	gggggaattccatgggtgtacagatagatatag, *Eco*RI [a] *Nco*I [b] ggcgggatcctcacttcacagccctagc, *Bam*HI [a,b]	pET302	BL21(DE3) Rosetta(DE3)
		Tfb4-8aa–AT-Fwr Tfb4-24a-Rev Tfb4-taatga-Rev	gggggaattcatatgagtgtacaaatagatatagtagg, *Eco*RI [a] *Nde*I [b] cggatccttcacagccctagcaactg, *Bam*HI [a,b] cggatcctcattacttcacagccctagcaactg, *Bam*HI [a,b]	pET11c, pET15b, pET24a	Rosetta(DE3) BL21(DE3)
		Tfb4-pQE30-Fwr Tfb4-pQE30-Rev	aaaggatccagtgtacagatagatatagtggg, *Bam*HI [b] aaaaagcttctacttcacagccctagcaac, *HindIII* [b]	pQE30	M-15
ttx_mbf1 (ttx_1938)	6xHis-*Ttx*-MBF1	1938-Rev 1938- p302-Fwr	aaaggatccctattcctcatcacgtagctcg, *Bam*HI [a,b] aaagaattccatggcacactactgcgacatatgcggc, *Eco*RI [a] *Nco*I [b]	pET302	Rosetta(DE3)
yeast_mbf1 (YOR 298C-A)	yMBF1	yMBF1fwr yMBF1rev	aaagaattcccttctcaaacacttcaacttaa, *Eco*RI [a,b] agcggccgcaatttattaccgagtcaaatata, *Not*I [a,b]	pRS316	DH5α

[a] Restriction enzyme used to clone into the subcloning vector, *pBlueScript* II KS(+) (ALTING-MEES & SHORT, 1989; ALTING-MEES et al., 1992)
[b] Restriction enzyme used to clone into the expression vector.
[c] Genes cloned by F. BLOMBACH, 2005
Recognition sites for restriction endonucleases are underlined.
For all genes the cloning vector *pBlueScript* II KS(+) was used. All oligonucleotides were purchased from INVITROGEN.

an initial denaturation step was performed at 95°C for 2 min and after completion of 30 cycles a final extension step was carried out at 72°C for 5 min.

2.4.7 Enzymatic manipulation of DNA.

2.4.7.1 Restriction of DNA.

Digestion of DNA was performed by incubating the double-stranded DNA molecules with an appropriate amount of restriction enzyme, in the respective buffer (1U of enzyme/ 1µg of DNA/ 50 µl total volume). The restriction reaction was performed using the optimal temperature and buffer recommended by the manufacturer (MEB Fermentas, St. Leon Rot, GER).

2.4.7.2 Ligation of vector and insert.

DNA ligation was performed by incubating the restricted DNA fragments with the restricted linearised vector in the presence of T4 DNA ligase (PAN et al., 1994; BAMKIER et al., 1987). DNA ligase catalyzes the formation of a phosphodiester bond between the 3' hydroxyl group of one nucleotide and the 5' phosphate group of another nucleotide. For ligation, restricted plasmid DNA (50 – 400 ng) and insert (1:1 up to 1:3) were mixed. 1 U T4 DNA ligase (MEB Fermentas, St. Leon Rot, GER) was added to a final volume of 20 µl. Ligation was carried out at 4°C overnight.

2.4.8 Transformation.

2.4.8.1 Preparation of chemically competent *E. coli* cells.

Competent cells from *E. coli* DH5α, BL21(DE3), BL21-CodonPlus(DE3)-RIL and Rosetta(DE3) strains were prepared by using the modified rubidium chloride/calcium chloride method (Promega Technical Manual, 1994); potassium chloride was introduced in the buffers instead of rubidium chloride.

Briefly, 100 ml LB-medium was inoculated with 1 ml overnight culture and incubated at 37°C in a rotary shaker until OD_{578nm} reached 0.6. The cell suspension was incubated for 5 min on ice and then centrifuged (6,000 x g, 6 min, 4°C). All following procedures were carried out on ice. The cell pellet was gently resuspended in 40 ml ice cold buffer 1 (30mM KAc, 10mM $CaCl_2$, 50 mM $MnCl_2$, 100 mM KCl, 15% glycerol) and incubated on ice for 5 min. After centrifugation (6,000 x g, 6 min, 4°C), the cell pellet was resuspended in 4 ml ice cold buffer 2 (10 mM MOPS, 75 mM $CaCl_2$, 10 mM KCl, 15% glycerol, pH 7.0) and aliquots (80 µl) were stored on -80°C for long term conservation.

2.4.8.2 Transformation of competent *E. coli* cells.

80 µl competent *E. coli* cells were gently mixed with plasmid DNA (1 – 3 µl) or ligation mixture (10 – 20 µl) and incubated on ice for 30 min. Transformation was achieved by incubation at 42°C for 45 sec. After the cells were stored for 5min on ice and the cells were transferred to 800 µl LB-medium and incubated for about 1 h at 37°C in a rotary shaker. 100 µl of transformed cells were plated on LB agar plates containing the respective antibiotic(s). The remaining 850 µl were centrifuged, the supernatant discarded and the pelleted cells were resuspended in about 100 µl LB-medium and also plated. After incubation of LB-agar plates at 37°C overnight, colonies were screened for positive clones carrying the recombinant plasmid DNA by boiling PCR and restriction analysis of the isolated plasmid DNA.

2.4.8.3 Preparation of electrocompetent *E. coli* cells.

100 ml LB-medium was inoculated with 1 ml overnight culture and incubated at 37°C in a rotary shaker until OD_{578nm} reached 0.3. The culture was centrifuged (6,000 x g, 15 min, 4°C) and the cell pellet was resuspended with 50 ml ice cold 10% glycerol. This step was repeated two times and the cell pellet was resuspended in 500 µl ice-cold 10% glycerol and aliquots (40 µl) were stored on -80°C for long term conservation.

2.4.8.4 Electroporation of electrocompetent *E. coli* cells.

40 µl of electrocompetent cells were defrozen on ice and 3 µl of plasmid DNA was added. Plasmid DNA prepared using Winston solution (see 2.8.4) was dialysed for 30 min in distilled water before electroporation. The mixture was transformed into a pre-cooled electroporation cuvette and incubated 30 sec on ice. The electroporation was achieved at 1.8 kV using cuvettes. After electrophoresis 1 ml SOC medium (2% (w/v) Bacto-tryptone 0.5% (w/v) Yeast Extract, 0.05% (w/v) NaCl, 2.5 mM KCl, pH 7.0, 10 mM $MgCl_2$; after autoclaving 20 mM Glucose) was added to the cuvette, the cells were incubated at 37°C for 45 min and plated on LB-agar plates containing the appropiate antibiotic.

2.4.9 Molecular cloning of TFBs from *T. tenax*.

The encoding genes *tfb2* (transcription factor B-2, TTX_2085), *tfb3* (transcription factor B-3, TTX_1929), *tfb4* (transcription factor B-4, TTX_1732), and *mbf1* (multiprotein bridging factor 1, TTX_1938) from *T. tenax* were cloned into the subcloning vector, *pBlueScript* II KS(+). Afterwards the inserted gene was excised from the subcloning vector and inserted into the pET vector system using the restriction sites introduced by PCR mutagenesis. The

expression vectors for cloning and expression hosts used in this study are shown in Table 1. The sequences of the cloned genes were verified by dideoxy sequencing (AGOWA, Berlin, GER) and expression of the recombinant enzymes in expression cells was performed as described in 2.5.1.

2.4.10 Molecular cloning of wild-type *T. tenax* promoter regions and mutated *fba-pfp* promoter fragments.

Cloning of the promoter regions of *tfb1, gar1-tfb2, fba-pfp, pps* and *orf1155-lrp* from *T. tenax* and the construction and molecular cloning of the mutated *fba-pfp* promoter fragments were accomplished by S. KONING in the laboratory of Prof. Dr. SIEBERS. The primers used for cloning of the *T. tenax* promoter fragments are indicated in Table 2. In all cases, the amplified fragments were cloned into *pBlueScript* II KS(+) via the restriction sites *Eco*RI and *Bam*HI. The mutated *fba-pfp* promoter fragments were constructed via PCR mutagenesis (Table 2) by S. KONING.

TABLE 2. List of primer sets for cloning of promoter regions and mutated *fba-pfp* promoter sequence in *E. coli*

Promoter region (gen number)	Primer Name	Sequence	Restriction enzyme[a]
*gar1-tfb2** (ttx_2086-ttx_2085)	Gar1F	ggggaattcgggcgtacgtcgtggatccgc	*Eco*RI
	Gar1R	ccgccggatcccccgattacgtcatagaaagt	*Bam*HI
*tfb1** (ttx_1484)	TFB1promF	ggggaattcggggaataaactaaataactaatatatgcg	*Eco*RI
	TFB1promR	ccgccggatcctggttagtgccctgagccgacatcc	*Bam*HI
*orf1155-lrp** (ttx_1154)	1154F-prom	cccccgaattccgacaccttgggcagtttgacc	*Eco*RI
	1154R-prom	ccgccggatcccgagcttctttattctggacg	*Bam*HI
*pps** (ttx_0910)	0910F-prom	cccccgaattcgatcctgaaactttagag	*Eco*RI
	0910R-prom	ccgccggatcccggacgagtctgg	*Bam*HI
fba-pfp (ttx_1278)	1278F-prom	cccccgaattcacgcgcatccaagag	*Eco*RI
	1278R-prom	ccgccggatcctgctcaatgccg	*Bam*HI
fba-pfpB/T (ttx_1278)	FBPA-TATA_R	cttctcggcgatgctgtctaaagtatattattaagtag	-
	FBPA-TATA_F	gcatcgccgagaagtggataattgctc	-
fba-pfpP (ttx_1278)	FBPA-1F	ggggaattcccgcctcaatggc	*Eco*RI
	FBPA-1R	gtagagcaattaatataatgg	-
	FBPA-2F	tattaattgctctacctcggcgcagatgctatg	-
	FBPA-2R	cgaatatccttaagaatgccggccactttgacg	-
	FBPA-3F	attcttaaggatattcg	-
	FBPA-3R	ggcggatccagtccgcgggcccgtgc	*Bam*HI

[a] Restriction enzyme used to clone into the subcloning vector, *pBlueScript* II KS(+)
Recognition sites for restriction endonucleases are underlined.
For all promoter regions the cloning vector *pBlueScript* II KS(+) was used.
(*) Cloning procedure performed previously by S. KONING in the laboratory of Prof. Dr. SIEBERS.
All oligonucleotides were purchased from INVITROGEN.

2.4.11 Automated DNA sequencing.

Automated DNA sequencing (SANGER et al., 1977) was performed by AGOWA (Berlin, GER). 10 µl of plasmid DNA (80 ng/µl) and, if required, 2 µl of specific primer (10 µM) were used for one sequencing reaction.

2.4.12 Labeling of DNA probes.

2.4.12.1 Oligonucleotide end-labeling with ^{32}P-γ-ATP.

10 µM of the oligonucleotide was mixed with 10 µCi of ^{32}P-γ-ATP and 1 U of T4 polynucleotide kinase (PNK) (MEB Fermentas, St. Leon Rot, GER). The mixture was incubated for 30 min at 37°C. Afterwards PNK was inactivated for 5 min at 95°C. 5 µM 5´-labeled reverse oligonucleotide (M13R) and 10 µM of unlabeled forward oligonucleotide (M13F) were used as PCR primers in order to amplify and label the DNA probe for Exo III footprinting assays.

2.4.12.2 ^{32}P-γ-ATP 5´-end-labeling of double stranded DNA probes.

Appropriate PCR products were radioactively labeled with 10 µCi ^{32}P-γ-ATP using 1 U of PNK. The reaction was accomplished for 20 min at 37°C and afterwards PNK was inactivated for 10 min at 65°C. The unincorporated ^{32}P-γ-ATP was removed using the PCR-purification kit from Qiagen (Hilden, GER).

2.4.13 Oligonucleotide hybridisation to construct 60 bp-length double strand DNA (*fba-pfp*WT-60bp).

The 60bp- length DNA fragment, *fba-pfp*WT-60bp, encompasses the region from -50 to +10 of the *fba-pfp* promoter. To construct *fba-pfp*WT-60bp, complementary oligonucleotides (3,000 pmol), *fba-pfp*WT-60bpF and *fba-pfp*WT-60bpR were dissolved in hybridisation buffer (10 mM Tris-HCl, pH 8.0, 0.1 mM EDTA, 0.2 M NaCl). The mixture was incubated for 5 min at 95°C in a water bath and was allowed to cool down to room temperature overnight. Double strand DNA (dsDNA) with the desired size was excised from a native 15% polyacrylamide gel. The gel slides were immersed in gel elution buffer (1 mM Tris-HCl, pH 7.0, 0.5 M NH_4Ac, pH 7.0) and incubated overnight at 37°C. DNA was purified using the PCR-purification kit from Promega (Mannheim, GER) and resuspended in TE buffer, pH 8.0.

2.5 BIOCHEMICAL METHODS

2.5.1 Heterologous expression of recombinant proteins in *E. coli*.

For protein expression prewarmed medium containing the appropriate antibiotic was inoculated with 2% (v/v) and incubated at 37°C in a rotary shaker. Protein expression was induced at OD_{600nm} = 0.6 – 0.8 by the addition of 1 mM Isopropyl-beta-D-thiogalactopyranoside (IPTG; Gerbu Biotech GmbH, Gaiberg, GER) and incubated for 3 – 4 hours. Afterwards the culture was chilled on ice and cells were harvested by centrifugation (6,000 x g, 15 min, 4°C) and finally stored at -80°C.

2.5.1.1 Expression and purification of *Ttx*-TFB1.

The expression and purification of *Ttx*-TFB1 from inclusion bodies was performed by F. BLOMBACH (2005). Briefly, the *ttx-tfb1* gene was cloned in pET11c and then expressed heterologously in *E. coli* BL21-CodonPlus(DE3)-RIL as described in 2.5.1. 3.8 g cells (wet weight) were resuspended in 20 ml buffer A (100 mM bis-Tris propane, pH 6.8, 100 mM NaCl, 5 mM EDTA, 10 mM DTT) containing 0.5 % (v /v) Triton-X100. Cells were disrupted by three fold passage through a French Press cell at 1,100 psi and centrifuged for 10 min at 5,500 x g and 4°C. The pellet was resuspended in 20 ml buffer A containing 0,5 % (v /v) Triton-X100, 2 mg DNase I, 4 mg lysozyme. After centrifugation (10 min, 5,500 g, 4°C), the pellet was resuspended in 20 ml buffer A containing 5 mM EDTA, 5 mM sodium-desoxycholate and centrifuged again. Following resuspension of the pellet in 30 ml buffer B (8 M urea, 100 mM bis-tris propane, pH 7.3, 100 mM NaCl, 60 mM DTT) the protein solution was diluted in buffer B to a final volume of 260 ml and inclusion bodies were allowed to solubilize for 24 h. The protein solution was transferred into an anaerobic chamber and dialysed against 1.8 L degassed buffer C (100 mM Tris, pH 7.5, 150 mM NaCl, 100 µM ZnCl, 1 mM DTT) to decrease urea concentration and allow protein refolding. After dialysis for 12 h the buffer was renewed and dialysis continued for another 12 h. This step was repeated twice and the final dialysis step was performed for 6 h, resulting in a final theoretical urea concentration of about 2 mM. Precipitated protein was removed by centrifugation for 20 min at 20,000 x g, 20° C. Heat precipitation (120 ml protein solution) was performed at 70° C for 15 min, followed by centrifugation (30 min, 20,000 x g, 4° C) for the removal of precipitated protein. The supernatant was loaded onto a heparin sepharose column (2.3 ml, 0.5 ml/min, Ø 10 mm, Amersham Parmacia, Uppsala) equilibrated in buffer C. After extended washing with seven volumes buffer C, protein was eluted using a two step gradient in buffer C (350 mM, 800 mM NaCl). The major part of *Ttx*-TFB1 eluted during the second step and

the respective fraction was collected. Finally, after addition of 10% (v/v) glycerol *Ttx*-TFB1 (230 µg/ml) was stored at -80° C.

2.5.1.2 Expression and purification of *Ttx*-TFB2.

The expression and purification of *Ttx*-TFB2 was performed by F. BLOMBACH (2005). Briefly, the *ttx-tfb2* gene was cloned into pET11c and then transformed and expressed heterologously in *E. coli* strain BL21-CodonPlus(DE3)-RIL as described in 2.5.1. 5 g cells (wet weight) were resuspended in 15 ml buffer C without DTT (-DTT) and cells were disrupted by three fold passage through a French Press cell (1,200 psi). After centrifugation (100,000 x g, 45 min, 4°C), the crude extract was diluted with two volumes of buffer C (-DTT) and incubated at 70°C for 15 min, followed by centrifugation (20,000 x g, 30 min, 4°C) to remove denatured proteins. Up to this point, no reducing agent was added because of decreased thermostability of *Ttx*-TFB2 in the presence of DTT in the heat precipitation step. After heat precipitation *Ttx*-TFB2 was dialysed against 1 L of buffer C with additional 10 mM DTT and loaded onto a heparin sepharose column (2.3 ml, 0.5 ml/min, Ø 10 mm, Amersham Parmacia, Uppsala) equilibrated in the same buffer. After washing the column with seven volumes of buffer C, protein was eluted by a two step gradient (350 mM, 800 mM NaCl) in buffer C. *Ttx*-TFB2 eluted at 800 mM NaCl and the respective fractions (0.75 ml) were collected. After addition of 10 % (v/v) glycerol *Ttx*-TFB2 (900 µg/ml) was stored at -80°C.

2.5.1.3 Expression and purification of *Ttx*-TBP.

The expression and purification of *Ttx*-TBP was performed by F. BLOMBACH (2005). The *ttx-tbp* gene was cloned in pET11c and the protein was expressed heterologously in *E. coli* BL21(DE3) as described in 2.5.1. 5 g cells (wet weight) were resuspended in 30 ml buffer D (50 mM Tris, pH 8.0, 50 mM NaCl) and disrupted by three fold passage through a French Press cell at 1,100 psi. After centrifugation (10,000 x g, 30 min) crude extract was diluted with one volume of buffer D. Incubation at 90°C for 15 min and subsequent centrifugation (20,000 x g, 30 min, 4°C) removed the majority of *E.coli* host proteins. 50 ml protein solution was dialyzed against 1 L of buffer D at 4°C overnight and subsequently loaded onto a Q Sepharose column (150 ml, 1 ml/min, Ø 25 mm, Amersham Pharmacia, Uppsala) equilibrated in the same buffer. The column was washed with 10 volumes of buffer D and *Ttx*-TBP was eluted with linear NaCl gradient (50 mM - 1 M) in buffer D. TBP eluted at approximately 600 mM NaCl and the fractions (10 ml) were collected. These fractions were analysed by SDS-PAGE and the fractions harboring the purified protein were combined and after addition of 10% (v/v) glycerol 1.3 mg/ml *Ttx*-TBP was stored at –80°C.

2.5.1.4 Expression and purification of 6xHis- *Ttx*-TFB2.

The *ttx-tfb2* gene was cloned in pET302 in order to add a N-terminal histidine tag to the protein. The plasmid was used to transform *E. coli* Rosetta(DE3) and the recombinant protein was expressed as described in 2.5.1. 10 g cells (wet weight) were resuspended in 15 ml buffer E (50 mM Na_2HPO_4/NaH_2PO_4, pH 8.5, 300 mM NaCl) and cells were disrupted by three fold passage through a French Press cell (1,200 psi). After centrifugation (10,000 x g, 45 min, 4°C), crude extract was diluted with two volumes of buffer E and incubated at 70°C for 15 min, followed by centrifugation (20,000 x g, 30 min, 4°C) to remove denatured proteins. Up to this point, no reducing agent was added because of the previously observed decreased thermostability of *Ttx*-TFB2 in the presence of DTT in the heat precipitation step (BLOMBACH, 2005). After heat precipitation 6xHis-*Ttx*-TFB2 was dialyzed against 1 L buffer E with 10 mM DTT at 4°C overnight and afterwards loaded onto a Ni-NTA column (1 ml, 0.5 ml/min, Ø 10 mm, Bio-Rad (München, GER)) pre-equilibrated in buffer E with 10 mM DTT. After washing the column with 5 volumes buffer E supplemented with 10 mM DTT and elution with 250 mM imidazole in buffer E, the respective fractions (2 ml) were collected. Fractions were analysed by SDS-PAGE and the fractions harboring the enriched protein were combined. Fractions were analysed by SDS-PAGE and the fractions harboring the enriched protein were combined, dialized overnight against 1 L 50 mM Tris.HCl, pH 8.0, 150 mM NaCl, 10 mM DTT and subjected to gel filtration on HiLoad 26/60 Superdex 200 prep grade pre-equilibrated in the respective buffer. Fractions were analysed by SDS-PAGE and the fractions harboring the purified protein were combined. After addition of 20% (v/v) glycerol 10 ml of 0.9 mg/ml 6xHis-*Ttx*-TFB2 was stored at -80°C.

2.5.1.5 Expression and purification of 6xHis-*Ttx*-TFB3.

The *ttx-tfb3* gene was cloned in pET15b in order to add N-terminal His tag to the protein. The recombinant plasmid was used to transform *E. coli* Rosetta(DE3) and the recombinant protein was expressed as described in 2.5.1. 10 g cells (wet weight) were resuspended in 15 ml buffer F (50 mM Na_2HPO_4/NaH_2PO_4, pH 8.5, 300 mM NaCl, 20% glycerol) and cells were disrupted by three fold passage through a French Press cell (1,200 psi). After centrifugation (100,000 x g, 45 min, 4°C), crude extract was diluted with two volumes of buffer F and incubated at 70°C for 15 min, followed by centrifugation (20,000 x g, 30 min, 4°C) to remove denatured proteins. After heat precipitation 6xHis-*Ttx*-TFB3 was dialyzed against buffer E with 10 mM DTT and 20% glycerol at 4°C overnight and afterwards loaded onto a Ni-NTA column (1 ml, 0.5 ml/min, Ø 10 mm, Bio-Rad (München, GER)) pre-equilibrated in buffer F

with 10 mM DTT. After washing the column with 5 volumes buffer F supplemented with 10 mM DTT and elution with 250 mM imidazole in buffer F, the respective fractions (2 ml) were collected. Fractions were analysed by SDS-PAGE and the fractions harboring the enriched protein were combined. After addition of 20% (v/v) glycerol 2 ml of 0.1 mg/ml 6xHis-*Ttx*-TFB3 was stored at -80°C.

2.5.1.6 Expression and purification of 6xHis-*Ttx*-MBF1.

The *ttx-mbf1* gene was cloned in pET302 to add an N-terminal His tag to the protein. The recombinant plasmid was used to transform *E. coli* Rosetta(DE3) and the recombinant protein was expressed as described in 2.5.1. 10 g cells (wet weight) were resuspended in 15 ml buffer E (50 mM Na_2HPO_4/NaH_2PO_4, pH 8.5, 300 mM NaCl) and cells were disrupted by three fold passage through a French Press cell (1,200 psi). After centrifugation (100,000 x g, 45 min, 4°C), crude extract was diluted with one volumes of buffer E and incubated at 70°C for 15 min, followed by centrifugation (20,000 x g, 30 min, 4°C) to remove denatured proteins. After heat precipitation 6xHis-*Ttx*-MBF1 was dialyzed against 1 L buffer E with 10 mM DTT at 4°C overnight and afterwards loaded onto a Ni-NTA column (1 ml, 0.5 ml/min, Ø 10 mm, Bio-Rad (München, GER)) pre-equilibrated in buffer E with 10 mM DTT. After washing the column with 5 volumes of buffer E supplemented with 10 mM DTT and elution with 250 mM imidazole in buffer E, the respective fractions (2 ml) were collected. Fractions were analysed by SDS-PAGE and the fractions harboring the enriched protein were combined, dialized overnight against 1 L 50 mM Tris.HCl, pH 8.0, 150 mM NaCl, 10 mM DTT and subjected to gel filtration on HiLoad 26/60 Superdex 200 prep grade pre-equilibrated in the respective buffer. Fractions were analysed by SDS-PAGE and the fractions harboring the purified protein were combined. After addition of 20% (v/v) glycerol 15 ml of 0.25 mg/ml 6xHis-*Ttx*-MBF1 was stored at -80°C.

2.5.2 Analytical protein methods.

2.5.2.1 Protein quantification.

Protein concentration was determined using the Bio-Rad Protein Assay based on the method of Bradford (BRADFORD, 1976), following the manufacturer's instructions with bovine serum albumin (BSA) as standard (2 – 10 µg/ml).

2.5.2.2 SDS-polyacrylamide gel electrophoresis (SDS-PAGE).

For protein analysis and separation, SDS polyacrylamide gel electrophoresis was used according to Laemmli (LAEMMLI, 1970). For proteins with a molecular weight range between

20 – 300 kDa, 12 – 15 % SDS-polyacrylamide gels were used. The resolving gel comprises 10% (w/v) acrylamide-biscarylamide (30%), 375 mM Tris (pH 8.8, RT), 0.1% (w/v) SDS, 0.67% (w/v) APS (100 mg/ml), 0.067% (v/v) TEMED and the stacking gel 4.0% (w/v) acrylamide-biscarylamide (30%), 125 mM Tris-HCl (pH 6.8, RT), 0,1% (w/v) SDS, 0.45% (w/v) APS (100 mg/ml), 0.15% (v/v) TEMED. First the resolving gel ingredients were mixed, poured between the two glass plates, covered with isopropanol and then allowed to polymerize for 20 min. After isopropanol was removed, the gel surfaces were washed with A. bidest and dried with a Whatmann paper before pouring the stacking gel solution. Directly after pouring the stacking gel, a comb was introduced between the two glass plates. The comb was removed after polymerization for approximately 20 min and gels were stored at 4°C before usage.

For preparation of protein samples from *E. coli* cells the cell pellet from 0.5 ml of a culture (10^9 *E. coli* cells; OD 600nm = 1) was dissolved in 50 µl distilled water, was mixed with 2 – 5 times concentrated loading buffer (final concentration: 62.5 mM Tris.HCl, pH 6.8, 10% glycerin, 2% SDS, 5% ß-mercaptoethanol, 0.005% bromophenol blue) and then heated for 5 min at 95°C. For preparation of protein samples from yeast cells, 1.5 ml of subclone cultures were transferred in an Eppendorf tube, the cells were spun down, 30 µl of glass beads and 100 µl of 1 time concentrated loading buffer were added and the sample was mixed using vortex for 20 sec. This step was repeated three times storing the tubes on ice in between and afterwards the protein sample was heated for 5 min at 95°C.

Electrophoresis was performed using a Minigel-Twin-Chamber (Biometra) at 20 mA. The electrophoresis buffer consisted of 25 mM Tris-HCl, 190 mM glycine and 0.1% v/v SDS (pH 8.3). As standard, PageRuler Marker (MEB Fermentas, St. Leon Rot, GER) was used with a molecular weight range of 14.2 – 66.0 kDa. Proteins were visualized by gel staining (40% methanol, 10% acetic acid and 0.25% Coomassie brilliant blue R-250) and destaining (Aqua bidest, cook the gel two times for about 5 min using the microwave). SDS gels were analysed and documented using the Gel Doc System in combination with the Quantity One Software Package (BioRad, München, GER).

2.5.2.3 Generation of antibodies using purified protein.

Polyclonal antibodies specific for 6xHis-*Ttx*-MBF1, 6xHis-*Ttx*-TFB2 and *Ttx*-TBP were generated commercially by EUROGENTEC (Seraing, BEL). SDS-PAGE and western blotting procedures were performed according to standard protocols.

2.5.2.4 Western blotting analysis.

Electrophoretically separated proteins (see 2.5.2.2) were transferred from the SDS-polyacrylamide gel to a Polyvinylidene Difluoride (PVDF) membrane (Roth). Therefore, the polyacrylamide gel was equilibrated in transfer buffer (50 mM Tris, 380 mM Glycine, 0.1% SDS, 20% methanol) for 15 min at RT. The membrane was then soaked together with two sponges and four Whatman filter papers (8 x 6 cm) in transfer buffer for 15 min at RT and the sandwich was assembled as follows, one sponge, two of the four pieces of filter papers, gel, nylon membrane, two filter papers and one sponge. The sandwich was transferred into the blotting chamber with the gel towards the anode. The protein transfer was performed overnight at 4°C applying an electric field (setting: 12V, 500 mA and 150W). After transfer the blot was air dried and then wetted in 20% methanol for 5 min to remove excess of SDS. Afterwards the membrane was washed three times for 5 min with buffer PBS-Tween (10x: 63.2 mM Na_2HPO_4, 11.7 mM KH_2PO_4, 68 mM NaCl, pH 7.4, after sterilizing by autoclaving 0,3% Tween-20 was added). The blocking step was performed for 1 h incubation of the membrane with 5% Blocking Buffer (5% dry-milk powder dissolved in buffer PBS-Tween). The primary antibodies were dissolved in 2.5% Blocking Buffer (2.5% dry-milk powder dissolved in buffer PBS-Tween). The blocking solution was removed and the membrane was incubated with the primary antibody at room temperature for 90 min (or overnight at 4°C) using a laminar flow shaker. After washing three times for 15 min at RT, the membrane was incubated for 90 min at RT with the secondary antibody conjugated to alkaline phosphatase. The membrane was washed 15 min with buffer PBS-Tween. This step was repeated three times. The immunoreaction was performed by incubating the membrane with 1 ml of CDP-Star (Tropix) in 9 ml of pre-warmed distilled water. After 5 – 10 min at 37°C the detection solution was removed by washing the membrane with distilled water for 10 sec. The membrane was dried for some seconds and afterwards wrapped in plastic. The detection was achieved by chemiluminiscence using the VersaDoc gel documentation system.

2.6 RADIOACTIVE ELECTROPHORESIS MOBILITY SHIFT ASSAY.

Protein-DNA complexes were detected on the basis of alterations in electrophoretic mobility. Binding assays (10 µl final volume) contained 12,5 fmol of 5´-end-labeled DNA fragments in binding buffer (BSA 0.1 µg/µl, 5% glycerol, 4 mM $MgCl_2$, 1 mM DTT, 20 mM Tris-HCl, pH 8.0, 100 mM NaCl). Finally, the respective purified transcription factor(s) were added and the mixture was incubated at 70°C for 20 min. For competition experiments, the unlabeled competitor DNA was added to the binding buffer on ice, afterwards the labeled probe and

purified transcription factors were immediately added, and the mixture was incubated at 70°C for 20 min prior to electrophoresis. The amount of protein and competitor DNA added in the experiments is indicated in each figure in section RESULTS (see results).

DNA-protein complexes were analysed on native 4% polyacrylamide gels (0.5x TBE (45 mM Tris-HCl, 45 mM borate, 1 mM EDTA, pH 8.3), 4.0% (w/v) acrylamide-biscarylamide (30%), 0.45% (w/v) APS (100 mg/ml), 0.15% (v/v) TEMED). Loading dye (1x) from Qiagen was used as loading buffer. Electrophoresis was performed at 200 V for approximately 2.5 h using 0.5x TBE buffer as running buffer.

The detection of the labeled DNA probes was carried out by autoradiography. Therefore, the gels were wrapped in cellophane and exposed to a standard X-ray film (BioMax ML, 18 x 24 cm). Exposure was carried out at -80°C in a light-tight BioMax cassette equipped with a BioMax MS intensifying screen. The films were manually processed in the darkroom using X-ray developer and fixer solutions (LX24 and AL4).

2.7 EXONUCLEASE III FOOTPRINTING ASSAY.

To perform footprinting experiments, the DNA probes were labeled using ^{32}P-γ-5'-end labeled universal primer M13R and the coding strand (mRNA like) was labeled. The binding step was achieved as described for EMSAs (see 2.6). However, the binding buffer contained: 40 mM Na-HEPES, pH 7.3, 6 mM $MgCl_2$, 0.1 mM EDTA, pH 8.0, 250 mM KCl, 0.1 µg/µl BSA, 5 ng/µl λ *Hind*III, 10% glycerol. After incubation of protein-DNA complex at 70°C for 20 min, 100 U Exonuclease III (MEB Fermentas, St. Leon Rot, GER) was added to a final volumen of 15 µl and the reaction was incubated for 20 or 40 min at 37°C.

The reaction was stopped by addition of 5 µl of 80% formamide loading dye buffer (19,6 ml 98% (v/v) Formamid, 0.1% (w/v) Bromophenol blue, 0.1% (w/v) Xylen Xyanol, 0.4 ml 500 mM EDTA pH 8.0, afterwards the solution was filtered using a syringue and 0.45 µm filter and store at -20°C). Finally the sample was heated at 95°C for 5 min.

For electrophoresis the Sequi-Gen Nucleic acid Electrophoresis Cell System (format: 50cm x 21cm) (Bio-Rad) was used. The glass plates were carefully cleaned with mild-soap, 90% ethanol and repellent solution (Gerbu) was added to one of the two glass plates. The non-native gel solution (30 ml 6% PAA [75 ml acrylamide-biscarylamide (Rotiphorese 40), 210 g urea, 50 ml 10x TBE, filtered solutions], 200 µl 10% (w/v) APS, 20 µl TEMED) was always freshly prepared and poured into the previously cleaned and assembled glass plates with the aid of a syringue. The combs were introduced between the two glass plates and the gel was polymerized for at least 2 hours at RT. Prior to sample application the gel was runned until

the temperature indicator reached 50°C. Then 7 µl of samples were loaded and the electrophoresis was performed for about 2 hours at 55 W using one time concentrated TBE running buffer (90 mM Tris, 90 mM borate, 2 mM EDTA, pH 8.3). Electrophoresis was stopped when the Bromophenol blue dye reached the botton of the gel. The detection of the labeled DNA probes was carried out by autoradiography as described in 2.6.

2.8 WORKING WITH YEAST AS GENETIC MODEL.

2.8.1 Yeast strains.

Growth and manipulation of yeast was performed according to standard procedures. AEY3780 (MATα *leu2*Δ0, *met15*Δ0, *ura3*Δ0) is the parental wild-type strain of the derivatives used in this study. The generation of the deletion strains lacking *MBF1* gene (Δ*mbf1*) was done by replacing the entire gene with *LEU2*. AEY3780 and 45E11 Δ*mbf1* yeast strain (MATα *leu2*Δ0, *met15*Δ0, *ura3*Δ0, *mbf1*Δ::*LEU2*) were kindly provided by Prof. Dr. ANN EHRENHOFER-MURRAY, Faculty of Biology and Geography, University of Duisburg-Essen.

2.8.2 Preparation of competent yeast cells.

100 ml of YPD medium (1% yeast extract, 2% Bacto peptone, 2 Bacto agar; after the sterilization process using the autoclave 2% glucose is added) was inoculated with 1 ml of a preculture and the culture was grown at 30°C (200 rpm) until an OD 600nm of 0.6 is reached. The cells were centrifuged (4,000 x g, 5 min, 4°C) and then washed with sorbitol buffer (1M sorbitol, 10 mM Bicin, pH 8.35, 3% Ethylenglycol). After centrifugation for 10 min at 4,000 x g and 4°C, the cells were resuspended in 1 ml sorbitol buffer and 55 µl DMSO is added. The competent cells were stored in 100 µl aliquots at -80°C.

2.8.3 Transformation of yeast competent cells.

5 µl of pre-warmed salmon sperm DNA (10 mg/ml) and 1 µg of yeast plasmid were added to 100 µl of frozen competent yeast cells and the mixture was incubated for 5 min at 37°C at 600 rpm and afterwards 700 µl of PEG/Bicin solution (40% PEG 1000, 200 mM Bicin, pH 8.35) was added. After 1 h of incubation at 30°C under static conditions, the suspension was centrifuged at 13,000 x g for 10 sec at RT in order to collet the cells in the botton of the tube. This step was repeated 6 times. After the final centrifugation 1 ml of NaCl/Bicin solution (150 mM NaCl, 10 mM Bicin) was added to the cell pellet and 900 µl of the supernatant was removed and the pellet was dissolved in the residual 100 µl. 20 µl of transformed cells were

plated on YM agar plates (2% (w/v) Bacto agar, after autoclaving, 2% (w/v) sterile glucose and 1x Yeast Nitrogen Base was added) supplemented with leucine and metionine. The remaining 80 µl were plated on an additional selective plate. Afterwards the YM-agar plates were incubated at 30°C for 2 to 3 days.

2.8.4 Preparation of genomic DNA from yeast using Winston solution.

20 – 30 ml YM-medium supplemented with 2% glucose and appropiate amino acids was inoculated with yeast cell suspension and incubated at 30°C overnight in a rotary shaker. The cell suspension was centrifuged (3,750 x g, 5 min, 4°C) and the cell pellet was resuspended in 500 µl distilled steril water and the cell suspension was centrifuged again (6,000 x g, 10 sec, RT) and the supernatant was removed. Afterwards 200 µl Winston solution (2% Triton X-100, 1% SDS, 100 mM NaCl, 10 mM Tris-HCl pH 8.0, 1mM EDTA), 0.3 g acid-washed glass pearls and 200 µl Phenol/Chloroform/Isoamylic alcohol (25:24:1) were added. After mixing with vortex at maximum speed for 5 min, 200 µl 1x TE buffer was added and the suspension was centrifuged (6,000 x g, 5 min, RT). The upper phase was transferred to a cleaned Eppendorf tube and 200 µl phenol/chloroform/isoamylic alcohol was added. The suspension was mixed at maximum speed for 3 min and then centrifuged for 2 min (6,000 x g, 5 min, RT). The upper phase was transferred to a cleaned Eppendorf tube and 1 ml of 100% Ethanol was added. The pellet was recovered after centrifugation (6,000 x g, 5 min, RT) and then washed two times with 70% ethanol. The pellet was dried and finally resuspended in 50 µl sterile distilled water.

2.8.5 Recombination/gap repair cloning technique for cloning of the *mbf1* gene from *T. tenax*, *M. mazei* and for the construction of chimeric genes.

A gap was created in the cloning vector by restriction digestion. Flanking sequences complementary to the end of the gapped vector were added to the gene of interest by using PCR. Afterwards, competent yeast cells were transformed using the gapped vector and insert, and the gap on the vector is repaired by homologous recombination inside the yeast cells (Figure 1).

To construct *pyMBF1*, the 1,667-Kb *Eco*RI/*Not*I genomic fragment encompassing the entire *MBF1* regulatory and coding region was cloned into the *pRS316* shuttle vector using the primers *yMBF1rev* and *yMBF1fwr* (Table 1). The plasmids encompassing the *MBF1* gene from *T. tenax* and *M. mazei* were generated by the recombination/gap repair cloning technique in yeast using *pyMBF1* as cloning vector (Table 3 and 4). Hereunto, the *MBF1* gene from *T.*

tenax and *M. mazei* were amplified by nested PCR using a 73 bp- length primer set. The 73 bp primer set (reverse and forward) encompassed 23 bp to the specific region complementary to the *MBF1* gene (reverse and forward) and 50 bp flanking sequences complementary to the ends of a gapped plasmid. The PCR product was purified after an agarose gel electrophoresis. The cloning vector was linearised by restriction digestion. The primer set, restriction enzyme and cloning vector used to clone *mbf1* gene from *T. tenax, M. mazei* and to construct chimeric genes are shown in Table 3 and Table 4.

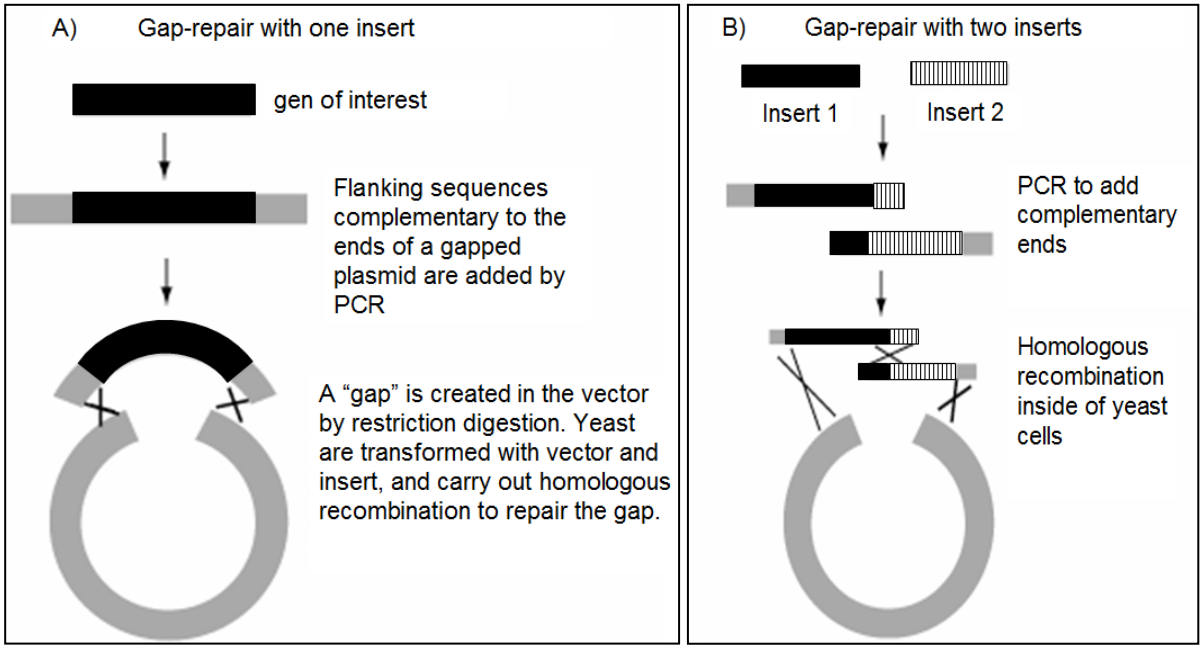


FIGURE 1. Basics of gap-repair cloning in yeast with one insert A) and two inserts B).

500 ng of the respective PCR product and 500 ng linear cloning vector were used to transform 45E11 Δ*mbf1* yeast strain and transformants were grown on YM agar plates supplemented with the amino acids leucine and methionine. The plates were incubated for 3 days at 30°C. The linear cloning vector was transformed also into 45E11 Δ*mbf1* yeast strain and used as control of the technique. Afterwards the transformants were recovered by dissolving the colonies from the plates in 1 ml of YM liquid medium, which was inoculated in 30 ml of YM liquid medium supplemented with leucine and methionine. After incubation overnight (200 rpm) at 30°C plasmid DNA was isolated using the Winston solution (see above 2.8.4). The plasmid DNA was used to transform *E. coli* DH5α by using electroporation (see above 2.4.8.4). After incubation of LB-agar plates at 37°C overnight, colonies were screened for positive clones carrying the recombinant plasmid DNA by boiling PCR, restriction analysis of the isolated plasmid DNA and by DNA sequencing (AGOWA).

TABLE 3. List of primers used for cloning y*MBF1*, T*MBF1*, M*MBF1* and chimeric genes to be expressed in yeast.

Primer Name	Sequence (5'-3')
yv-TNt-f	taaaagctaataagctagaaaacaaaaaagttaacgagcaaaatcgtaaagaaaaatgcactactgcgacatatgc
TCt-yv-r	tgcttcattgatgacatgcagtgcgaaaagaaaggaacaaatgaaagaagacctctctattcctcatcacgtagctc
yv-MNt-f	taaaagctaataagctagaaaacaaaaaagttaacgagcaaaatcgtaaagaaaaatgcagtgcgaaatatgtggt
MCt-yv-r	tgcttcattgatgacatgcagtgcgaaaagaaaggaacaaatgaaagaagacctctttacttgcgctttattttcac
yv-yNt-f	taaaagctaataagctagaaaacaaaaaagttaacgagcaaaatcgtaaagaaaaatgtctgactgggatacaaat
yv-THTH-yv-r	gaagacctcttcatttcttctttggagctcccaaaggcgaaccgatgttgttaccctcgagcaactttacgccaag
yv-MHTH-yv-r	gaagacctcttcatttcttctttggagctcccaaaggcgaaccgatgttgttaccttcggtaagtttaatgttaag
yHTH-Tv-r	aagctctcttctgttgtatcctccagactgatcctcggcctgttgaacgggtctcaatttaacgcctaaggc
yHTH-Mv-r	ggtggtgtctttcttcatatgagagacttcctgccccgcatcatccaagcgtctcaatttaacgcctaaggc
TNt-yv-r	tctgaccctcgttgtcacccctcgtattggttgatccgtacttcttgtcaacggaaacaacgccaccataagacctagc gcaacgc
yHTH-f	gttgtttccgttgacaagaagtacgg
MNt-yv-r	tctgaccctcgttgtcacccctcgtattggttgatccgtacttcttgtcaacggaaacaaccttaccgtacggtgcacatt tctggc
yHTH-yv-r	tgcttcattgatgacatgcagtgcgaaaagaaaggaacaaatgaaagaagacctcttcatctcaatttaacgcctaag g

TABLE 4. Plasmids generated by using the recombination/gap repair cloning technique in order to express yMBF1, TMBF1, MMBF1 and chimera in yeast.

Plasmid name	Primer sets used for cloning	Cloning vector (restriction enzyme) [a]
pTMBF1	yv-TNt-f, TCt-yv-r	pyMBF1 (*Box*I, *Eco*85I)
pMMBF1	yv-MNt-f, MCt-yv-r	pyMBF1 (*Box*I, *Eco*85I)
pTTyMBF1	yv-TNt-f, yv-THTH-yv-r	pyMBF1 (*Box*I, *Eco*85I)
pyTyMBF1	yv-yNt-f, yv-THTH-yv-r	pyMBF1 (*Box*I, *Eco*85I)
pMMyMBF1	yv-MNt-f, yv-MHTH-yv-r	pyMBF1 (*Box*I, *Eco*85I)
pyTTMBF1	yv-yNt-f, TCt-yv-r	pyMBF1 (*Box*I, *Eco*85I)
pyMyMBF1	yv-yNt-f, yv-MHTH-yv-r	pyMBF1 (*Box*I, *Eco*85I)
pyMMMBF1	yv-yNt-f, MCt-yv-r	pyMBF1 (*Box*I, *Eco*85I)
pTyyMBF1	yv-TNt-f, TNt-yv-r	pyMBF1 (*Bgl*II)
pyyTMBF1	yv-yNt-f, yHTH-Tv-r	pTMBF1 (*Van*96I)
pMyyMBF1	yv-MNt-f, MNt-yv-r	pyMBF1 (*Bgl*II)
pTyTMBF1	yv-TNt-f, TNt-yv-r; yHTH-f, yHTH-Tv-r	pTMBF1 (*Van*96I)
pyyMMBF1	yv-yNt-f, yHTH-Mv-r	pMMBF1 (*Bsi*WI)
pMyMMBF1	yv-MNt-f, MNt-yv-r; yHTH-f, yHTH-Mv-r	pMMBF1 (*Bsi*WI)
PyyΔCtMBF1	yv-yNt-f, yHTH-yv-r	pTMBF1 (*Van*96I)

[a] Restriction enzyme used to linearize the cloning vector

The yeast *MBF1* gene encoded by the vector *pyMBF1* was replaced by *MBF1* gene from *T. tenax* and *M. mazei* to construct plasmids *pTMBF1* and *pMMBF1*, respectively. *pTMBF1* and *pMMBF1* carry the entire yeast *MBF1* regulatory region and the respective *MBF1* coding region from *T. tenax* or *M. mazei*, respectively. *pTMBF1*, *pyMBF1* and *pMMBF1* were linearized and used as cloning vector to introduce *MBF1* chimeric genes by recombination/gap repair cloning technique (Table 3 and 4).

Yeast-archaeal MBF1 chimera were constructed by the use of the yeast recombination system. Chimeric proteins were named combining a three letter code: y, T or M, corresponding to the origin from yeast (y), *T. tenax* (T) or *M. mazei* (M) and the respective position at the N-terminal domain, core domain (flexible linker and helix-turn-helix- domain) or C- terminal part.

2.8.6 Complementation test: sensitivity to aminotriazole.

Yeast colonies of the wild-type (WT) strain and the *mbf1*Δ yeast strain transformed with plasmids expressing *T*MBF1, *M*MBF1, yMBF1 and chimeric proteins or empty vector were diluted in 1 ml of YM liquid medium. 200 µl of cell suspension was dissolved in 800 µl of distilled water, the optical density (OD_{600nm}) was determined and 200 µl of each cell suspension with OD_{600nm} of 0.3 (calculated using the equation, µl of cell suspension = 12 / OD_{600nm}) was transferred to the first row of the 96-well microtiter plate. 5-time serial dilutions of each cell suspension were performed by transferring aliquots (33.33 µl to 166.66 µl water) into the following six rows of the microtiter plate. A replica plater (6 x 8 pronged) was dipped into the microtiter plate to coat the ends of the prongs with the yeast suspensions and the replica plater was pressed gently onto YM plate supplemented with the amino acids, leucine and methionine, in the presence or absence of 3 mM aminotriazole. The plates were incubated for 3 days at 30°C.

2.9 SEQUENCE ANALYSIS.

The information derived from DNA sequencing was visualized by using CHROMAS software. BLAST searches were performed at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov, ALTSCHUL et al., 1997) and at the Integrated Microbial Genome (IMG) (http://www.img.jgi.doe.gov) (MARKOWITZ et al., 2009). Conserved genomic context analysis was performed via the IMG web site.

The CLUSTAL_X program (THOMPSON et al., 1994) with default settings was used to compare protein and DNA sequences. The aligned sequences were inspected and adjusted manually to minimize the number of gaps and insertions. These manual adjustments were based on the sequence similarities and secondary structure predictions. The secondary structure prediction was performed by the program HHpred (SÖDING, 1995) available on the web site Bioinformatics Toolkit at Max-Planck Institute for Developmental Biology, Tübingen (http://toolkit.tuebingen.mpg.de/hhpred).

The evolutionary inference of the TFBs and MBF1 proteins and their related sequences was performed according to the Neighbor-Joining method (SAITOU & NEI, 1987) and the phylogenetic tree was visualized by the MEGA program version 4.0 (TAMURA et al, 2007). The sampling variance of the distance values was estimated from 1,000 bootstrap resampling of the alignment columns.

3. RESULTS

Cells responding to dramatic environmental challenges or undergoing a developmental switch typically change the expression of numerous genes. In Bacteria, sigma factors help to regulate these responses at the level of basal transcription, whereas in Eukaryotes, four RNA polymerases and multiple general transcription factors (GTFs) are required. Transcription in the Archaea has elements resembling both Eukaryotes and Bacteria (BELL & JACKSON, 1998; GEIDUSCHEK & OUHAMMOUCH, 2005). However, relatively little is known, about how GTFs in Archaea help to confer fitness across a broad range of environments, including hostile ones, like thermal and hypersaline ponds, but also to milder environments, like oceans and human oral cavity and gut (DELONG & KARL, 2005; LEPP et al., 2004; VENTER et al., 2004). The archaeal basal transcription machinery is considered as a simpler model of the eukaryotic counterpart. Eukaryotes need at least six GTFs, whereas in Archaea, only two GTFs orthologous to eukaryotic transcription factor IIB (TFB) (COLANGELO et al., 2000) and TATA-binding protein (TBP) are necessary and sufficient for initiating basal transcription (GEIDUSCHEK & OUHAMMOUCH, 2005). The archaeal promoter architecture is similar to the one found in Eukaryotes, consisting of the three elements, TATA-box, BRE and Inr.

Interestingly, GTFs are present in multiple copies in several archaeal species. *Halobacterium* NRC-1 is particularly complex, because its genome encodes six TBPs and seven TFBs (BALIGA et al., 2000) and their function has been studied in detail. In contrast, crenarchaeal genomes encode generally one TBP and multiple homologues of TFB. All of them remain uncharacterized with the exception of the TFB1 homologues from *S. solfataricus* and *S. acidocaldarius* (BELL et al., 1999; 2000) and the more recently described TFB3 homologue from *S. solfataricus* (PAYTUBI & WHITE, 2009).

3.1 SEQUENCE AND GENE CONTEXT ANALYSIS OF TFBs IN *T. tenax*.

In the genome of *T. tenax* (SIEBERS et al., manuscript in preparation) four putative TFB homologues and one TBP homologue have been identified and designated as *Ttx*-TFB1 – 4 and *Ttx*-TBP, respectively. The aim of the current work includes the biochemical and functional characterization of TBP and TFBs homologues from *T. tenax*.

The single *T. tenax* TBP gene (*tbp*) codes for a protein of 203 amino acids with a calculated molecular weight of 22.1 kDa. *Ttx*-TBP shares 86% sequence similarity with the TBP homologue of *Pyrobaculum aerophilum* (genbank accession: AAL63996), and 56% and 57% aa identity to the characterized TBP homologues of *Sulfolobus shibatae* (genbank accession:

AAC43403) and *S. acidocaldarius* (genbank accession: CAA64405) (BELL & JACKSON, 2000; QURESHI et al., 1997), respectively.

BLASTP searches of all available databases including finished and unfinished genomes were conducted with the entire sequences of *Ttx*-TFB1 (323 aa), *Ttx*-TFB2 (300 aa), *Ttx*-TFB3 (160 aa) and *Ttx*-TFB4 (256 aa) from *T. tenax* as queries. The searches using *Ttx*-TFB1 and *Ttx*-TFB2 sequences identified homologues that exhibited at least 79% and 56% overall aa identity (E-value > 1e-50) to other TFB1 homologues within the domain Archaea and TFB2 homologues within the class Thermoprotei, respectively. For TFB3 and TFB4 no homologues with such high identity were identified in other archaeal genomes. For *Ttx*-TFB3, homologues were only identified only in the crenarchaeal phylum (Thermoproteaceae (41% identity, E-value > 1e-20) and Sulfolobaceae and Desulfurococcaceae (22% aa identity, E-value > 10e-0)). For TFB4 the blast searches revealed no significant hits (e.g. TFB1 homologues from members of the Thermoproteaceae, *I. hospitales*, *T. acidophilum*, Methanosarcinaceae, Methanobacteriaceae, Methanococcaceae, Methanopyraceae, Methanomicrobiaceae (24% aa identity, E-value < 1e-5) and TFB1 and TFB2 homologues from other Archaea (26% aa identity, E-value < 10e-0)). Therefore, these analyses suggest that TFB1 is highly conserved at the protein level within the archaeal domain whereas TFB2 and TFB3 are only conserved within the phylum Crenarchaeota and TFB4 is found in only one organism, *T. tenax*.

In order to analyze the phylogenetic relationship, a subset of 46 sequences of crenarchaeal TFBs was aligned (17 TFB1s, 15 TFB2s, 14 TFB3s, Figure 2 and 3). TFB4 was omitted since no obvious homologue was identified. Although the number of homologues is relatively small, three distinctive clusters can be observed corresponding to the three TFBs: TFB1, TFB2 and TFB3 (Figure 2). The basal relationships among the major branches are not supported by high bootstrap values (BV), but nevertheless there is strong support by the specification within each branch. The tree reflects several higher level taxonomic groups within each branch, for example, TFB1 homologues from Sulfolobales (TFB1, BV 100%; TFB2, BV 100% and TFB3, BV 65%) and from Thermoproteales (TFB1, BV 95% (except, *Thermophilum pendens*); TFB2, BV 100% (except, *T. pendens* and *Caldivirga maquilinguensis*); and TFB3, BV 99% (except one of the TFB3 homologues of *Pyrobaculum calidifontis*).

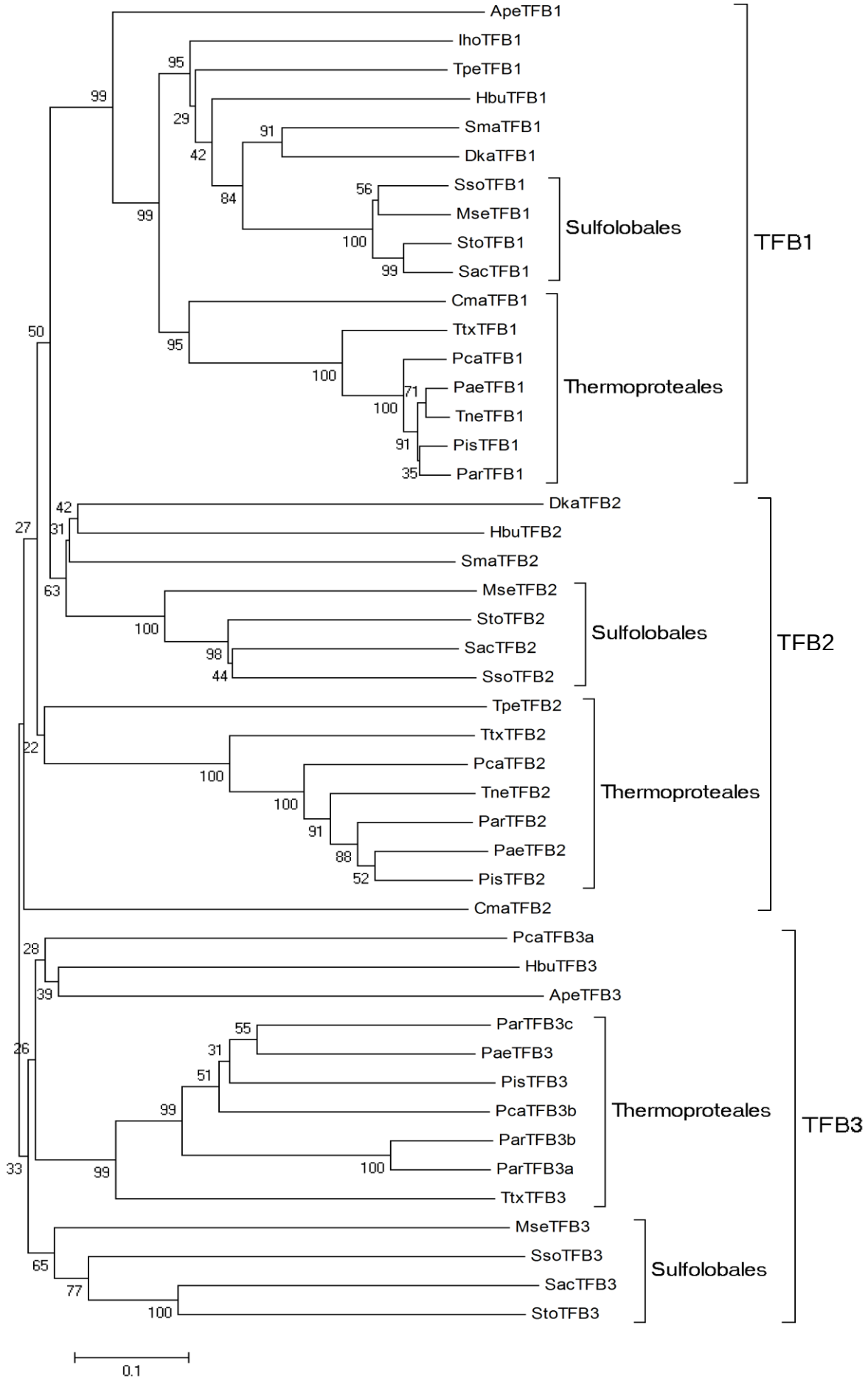
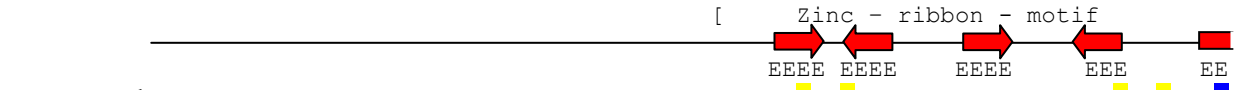

ApeTFB1
IhoTFB1
TpeTFB1
HbuTFB1
SmaTFB1
DkaTFB1
SsoTFB1
MseTFB1
StoTFB1
SacTFB1
Sulfolobales
CmaTFB1
TtxTFB1
PcaTFB1
PaeTFB1
TneTFB1
PisTFB1
ParTFB1
Thermoproteales
TFB1
DkaTFB2
HbuTFB2
SmaTFB2
MseTFB2
StoTFB2
SacTFB2
SsoTFB2
Sulfolobales
TpeTFB2
TtxTFB2
PcaTFB2
TneTFB2
ParTFB2
PaeTFB2
PisTFB2
Thermoproteales
CmaTFB2
TFB2
PcaTFB3a
HbuTFB3
ApeTFB3
ParTFB3c
PaeTFB3
PisTFB3
PcaTFB3b
ParTFB3b
ParTFB3a
TtxTFB3
Thermoproteales
MseTFB3
SsoTFB3
SacTFB3
StoTFB3
Sulfolobales
TFB3
0.1

FIGURE 2. Phylogenetic tree of 46 TFB homologues in Crenarchaeota.
The phylogenetic analysis is based on a multiple alignment of 46 crenarchaeal amino acid sequences (Figure 3) and was constructed using the Neighbor-Joining method (SAITOU & NEI, 1987). The bootstrap consensus tree inferred from 1,000 replicates (FELSENSTEIN, 1985) is taken to represent the evolutionary history of the taxa analyzed. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1,000 replicates) is shown next to the branches (FELSENSTEIN, 1985). All positions containing alignment gaps and missing data were eliminated in pairwise sequence comparisons (pairwise deletion option). Phylogenetic analyses were conducted in MEGA4 (TAMURA et al, 2007).

It is noteworthy that TFBs homologues from Desulfurococcales (i.e. *Aeropyrum pernix*, *Ignococcus hospitalis*, *Hyperthermus butylicus*, *Sthaphylotermus marinus*, *Desulfurococcus kamchatkensis*) are less conserved on protein level; none of the Desulfurococcales TFB groups in a cluster is supported by high bootstrap values (Figure 2). However, the phylogenetic tree nicely reflects the overall differentiation of subgroups of Thermoproteaceae, Sulfolobaceae and Desulfurococcaceae within each cluster.

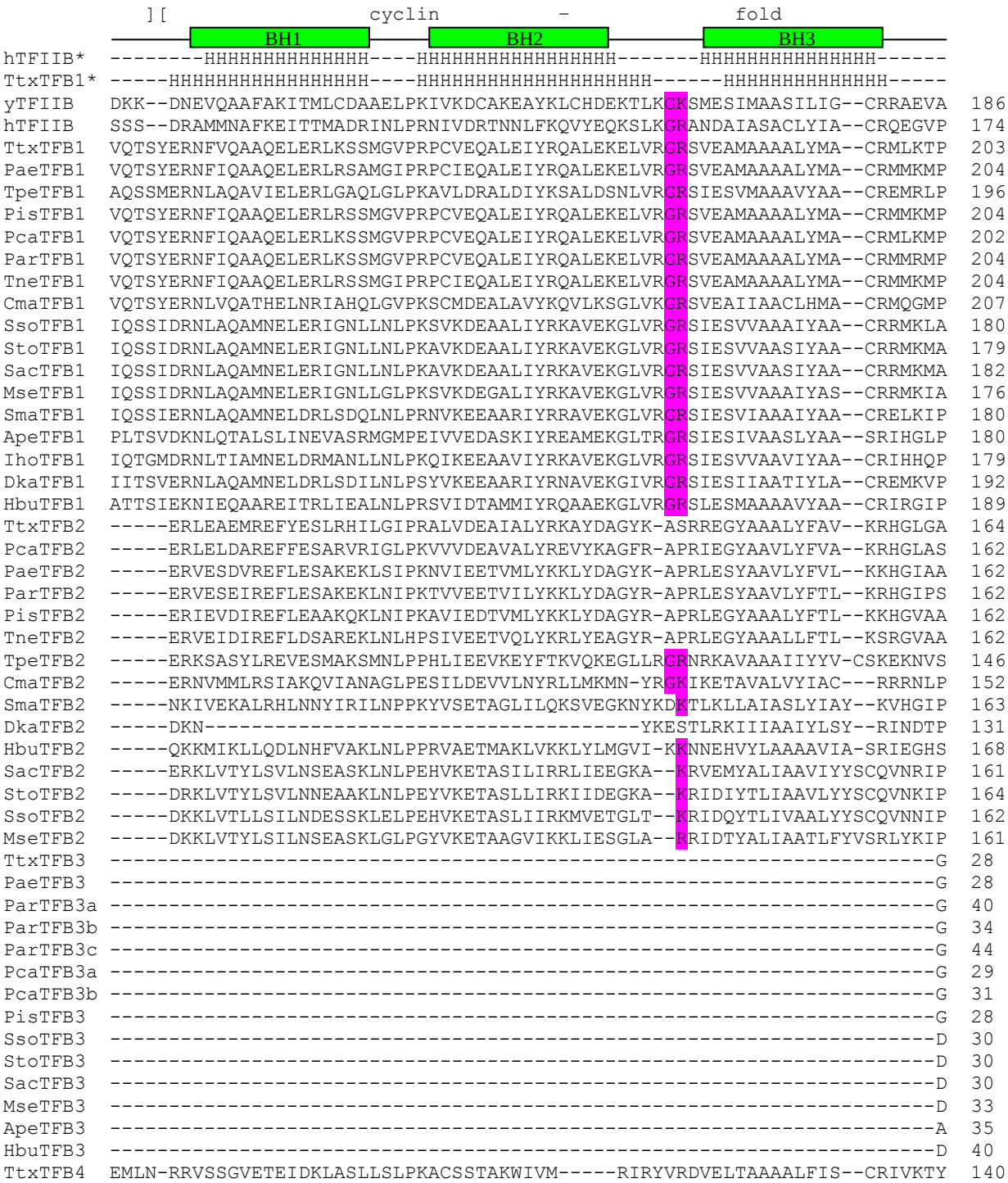
Remarkably, all crenarchaeal genomes encode at least two homologues of TFBs: TFB1 and TFB2. The third homologue, TFB3, which is not widely distributed in the phylum, is present in almost all Sulfolobales and Thermoproteales. Some members of the Thermoproteales even encode additional homologues of TFB3, for example, two in *Pyrobaculum calidifontis* and three in *Pyrobaculum arsenaticum*. Homology searches and phylogenetic analyses based on multiple sequence alignments indicated that TFB1 is more conserved within the group of Crenarchaeota than TFB2 and TFB3 (Figure 2 and 3) suggesting that gene duplication occurred at an early stage during the evolution of the phylum.

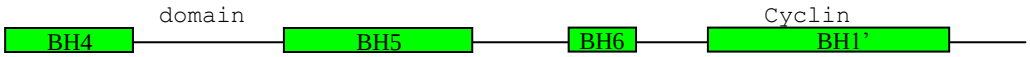
The crystal structure of TFB1 of the Euryarchaeon *Pyrococcus woesei* has been solved (KOSA et al., 1997). The pyrococcal TFB is an example of a classical TFIIB. The amino-terminal part of the *Pwo*-TFB comprises a zinc-ribbon structure, followed by a B-finger, which is important for RNA-polymerase interaction and two TFB-repeats (cyclin- fold domain). The TFB-repeats consist of five α-helices each and are essential for DNA- and TBP-interaction. A detailed structural inspection was performed on the basis of the TFB1-4 sequence alignment for crenarchaeal TFBs (Figure 3) using the conserved domain search tool (http://www.ncbi.nlm.nih.gov/Structure/cdd). The analysis reveals that the canonical Zn-ribbon motif is present in most TFB1-4 homologues, except TFB1 from Sulfolobaceae and Desulfurococcaceae (with the only exception of *I. hospitalis*) (Figure 3 and 4) and TFB2 from *D. kamchatkensis* and *Metallosphaera sedula*, which contain a non-canonical Zn-ribbon motif (Figure 3 and 4).



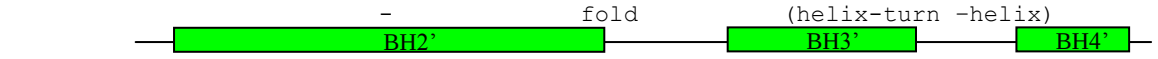

```
yTFIIB    1 ---------------------MMTRESIDKRAGRRGPNLNIVLTCPECKVYPPK-IVERFSEGDVVCALCGLVL 52
hTFIIB      ------------------------------MASTSRLDALPRVTCPNHPDAI---LVEDYRAGDMICPECGLVV 41
TtxTFB1   1 -------MS-AQGTNQGVPPKLRIAKSSDG-YLSLITETGEPYVCPVCGNDK---FVYNYERGEIVCVVCGTVI 62
PaeTFB1   1 -------MSVSNPPSSGKPLKLRVNKDSEG-YLSLVTESGEVYKCPICGNDK---FVYNYERGEAVCIVCGAVV 63
TpeTFB1   1 ---------------MGFGRESMVEERKEN-VNPEDSEEERELRCPNCGSTR---LIYDATRGEIICANCGYVM 55
PisTFB1   1 -------MSSSSPTSSGKPLKLRIDRDNEG-YLSLVTDTGEVYRCPICGNDR---FVYNYERGEIVCIVCGAVV 63
PcaTFB1   1 -------MSTATPP--GKPLKFRIHRDSEG-YLSLVSETGEVYSCPVCGNDK---FVYNYERGEVVCIVCGTVV 61
ParTFB1   1 -------MSSTSLPSSGKPLKLRINRDSEG-YLSLVTESGEIYRCPICGNDR---FVYNYERGEVVCIVCGAVV 63
TneTFB1   1 -------MTSPNPPSSGKPLKLRINRDSEG-YLSLVTDTGEVYRCPICGNDK---FIYNYERGEVVCIVCGAVV 63
CmaTFB1   1 MSTTQGGSNQQDETRRELIIQKFKLFRREGNNLEFVDQSGQKLRCPICGNTV---FIEDAERGQIVCASCGYVL 71
SsoTFB1   1 ----------------------------------MLYLSEENKSVSTPCPPDK---IIFDAERGEYICSETGEVL 38
StoTFB1   1 -----------------------------------MSESNKSQASATPCPPDK---IVFDEERGEYICTETGEVI 37
SacTFB1   1 --------------------------------MGGEMVEQSKVPSSSLCPPDK---IIFDEERGEYICTETGEVI 40
MseTFB1   1 ----------------------------------MSKEN---SSDTQCPPDK---IIFDADRGEYICSETGEVI 34
SmaTFB1   1 --------------------------------MSNPVN--TSPKDKQKCPEDQ---IIYDPEMGEYICLLTGEVI 38
ApeTFB1   1 ----------------------------MASEIPYDESPSGEESGEIKCKN-----IVTDPVRGLKICADTGEII 42
IhoTFB1   1 ----------------------------------MPEQEAFRLRCPVCGSTD---IVFNEETGEYVCARCGTIV 37
DkaTFB1   1 --------------------MGVLTMFIDSTASSPESSMDSESKKPKCPQDK---IVYDTTHGEYVCQDTGEVI 51
HbuTFB1   1 ----------------------MFEEGYTPVEGGEESEAKEGIGDRCPPEY---IVFDEERGEYICTLTGEVV 48
TtxTFB2   1 ------------------------------MTRRILADLAEQDLICPVCGAVNK--VKIDDERGSIVCTECGAVI 43
PcaTFB2   1 ------------------------------MTRRLIFEVEEY--RCPVCGSDKN--IVIDYEGGQIICKNCGTVL 41
PaeTFB2   1 ------------------------------MTRKLIFELEEY--ACPACGAVND--VIVDYERGQVICKSCGVVL 41
ParTFB2   1 ------------------------------MTRRLIFEIEEQ--VCPVCGSKGE--ITIDHERGQVVCRACGTVV 41
PisTFB2   1 ------------------------------MTRKLIFEVEEH--RCPVCGAVND--VVVDYERGQVICKNCGTVL 41
TneTFB2   1 ------------------------------MTRRLIFELEEY--RCPTCGAVND--VAVDYERGQVVCRSCGTVL 41
TpeTFB2   1 ------------------------------------------MSCPNCGRKVK--PILNFETGEYACPYCGYVF 30
CmaTFB2   1 ---------------------------------------MNKCPACGSSE---FIYDEYSGTVYCAKCGYVL 30
SmaTFB2   1 ------------------------------------MKYEHNQLVCPYCGSHS---VIFDYETNEYVCTRCGAVI 36
DkaTFB2   1 ----------------------------MVEEKLTYQGFEGFLKNNNTPRYSG---MFTNMVHDHGVGGTEISGY 44
HbuTFB2   1 ------------------------------------MSIPLPKGLEYCPVCKAPLKL-YAKRDGSK-LVCTNCGYVI 39
SacTFB2   1 ---------------------------------------MKCKICGSES---LIFDRERGIYVCINCASVD 29
StoTFB2   1 ----------------------------------------MEKIKCPNCGSES---ITLDLNKGTYICTNCGYVI 32
SsoTFB2   1 ---------------------------------------MKCPYCKTDNA--ITYDVEKGMYVCTNCASVI 30
MseTFB2   1 ---------------------------------------MKCPICEGTE---IIYDAEHGNYVCARDGTVI 29
TtxTFB3   1 ---------------------------------------MRCSYCGSDD-----IIISSGEYVCRKCGSVL 27
PaeTFB3   1 ---------------------------------------MKCPYCQSMN-----ITLADGEYVCRDCGTVL 27
ParTFB3a  1 ----------------------------MTIRGAARAAFPWPSCPHCQSAN------AAADGEYVCRECGTVI 39
ParTFB3b  1 ------------------------------------VWGAFPWPKCPFCQSAN------VVADGEYVCRECGTVI 33
ParTFB3c  1 -------------------------MKYTYYTKNQIKYLFVMICPYCKSTN-----ITLADGEYVCRDCGTVL 43
PcaTFB3a  1 ----------------------------------------MACPVCGG-R---AVVVSPDGEYVCVQCGTVI 28
PcaTFB3b  1 ---------------------------------------MRHMKCTYCNSDK-----VILVDGEYVCTECGTVV 20
PisTFB3   1 ---------------------------------------MICPYCRSDK-----IILKDGEYICTECGTVL 27
SsoTFB3   1 ---------------------------------------MQCPACGSNE---IIWDNKNGEIVCSNCGTII 29
StoTFB3   1 ---------------------------------------MKCPYCESEN---LVWDYKNGNLVCTSCGSVI 29
SacTFB3   1 ---------------------------------------MECPECKSRE---IVWDYKCGNLVCSNCGLVL 29
MseTFB3   1 ----------------------------------------MDNIRCSYCGSKN---LIWDYTRGEVVCGDCGTTA 32
ApeTFB3   1 ------------------------------------MDGRAGSCSYCGSTM---VVTLHTEGEIVCGSCGAVL 34
HbuTFB3   1 ----------------------------------MVLYYPLIRLEKCPYCGSSK---LIVESARGQLVCAACGAVL 39
TtxTFB4   1 ----------------------------------VSVQIDIVGRCPVCGGER---LALN--EYMVVCTSCGTVL 35
```


][B - finger - motif][linker
PfuTFB1
EE +
yTFIIB SDKLVDTRSEWRTFSNDDHNGDDPSRVGEASNPLLDGNNLSTRI------GKGETTDMRFTKELNKAQGKNVM 119
hTFIIB GDRVIDVGSEWRTFSND-KATKDPSRVGDSQNPLLSDGDLSTMI------GKGTGAASFDEFGNSKYQNRRTM 107
TtxTFB1 SENLVDLGPEWRAFTNEEKGQRARTGAPITR-LVSEALTTVIDWRDRDVSGKELDLKRKLEVIRLRKWQTRAR 134
PaeTFB1 QEQLLDLGPEWRAFTSEEKGQRARTGAPLTR-LISEALTTVIDWRDKDVSGKELDIKRKLEVIRLRKWQTRAR 135
TpeTFB1 SEREIDQGAEWRAFTPEEREKRSRVGAPISRYGVES-LVTDIDWGGRDAAGREISLRKRIEMLRLRKWQVRAR 127
PisTFB1 QEQLLDLGPEWRAFTSEEKGQRARTGAPLTR-LISEALTTVIDWRDKDVSGKELDIKRKLEVIRLRKWQTRAR 135
PcaTFB1 QEQLLDLGPEWRAFTSEEKGQRARTGAPLTR-LISEALTTVIDWRDRDVSGKELDIKRKLEVIRLRKWQTRAR 133
ParTFB1 QEQLLDLGPEWRAFTSEEKGQRARTGAPLTR-LISEALTTVIDWRDKDVSGRELDIKRKLEVIRLRKWQTRAR 135
TneTFB1 QEQLLDLGPEWRAFTSEERGQRARTGAPLTR-LISEALTTVIDWRDKDVSGKELDIKRKLEVIRLRKWQTRAR 135
CmaTFB1 MEHILDTGPEWRAFTPEEKEDRARTGGPLER-VTSEELVTRIETTLKSP-----DLKKKLEILKYKKWQQRIR 138
SsoTFB1 EDKIIDQGPEWRAFTPEEKEKRSRVGGPLNNTIHDRGLSTLIDWKDKDAMGRTLDPKRRLEALRWRKWQIRAR 111
StoTFB1 EERAIDQGPEWRAFTPEEKEKRSRVGGPLNQTIHDMGISTVIDWKDKDAMGRTLDPKRRLEVLRWRKWQIRAR 110
SacTFB1 EERIIDQGPEWRAFTPEEKEKRSRVGGPLNQTIHDMGISTVIDWKDKDAMGRSLDPKRRLEVLRWRKWQIRTR 113
MseTFB1 EERIVDQGPEWRAFTPEEKEKRSRVGGPLNQTIHDRGLSTLIDWKDKDAIGRNLDPKRRLEVLRWRKWQIRAR 107
SmaTFB1 EEKVIDDRPEWRAFTPEERERRSRIGGPLTSTVHDRGLATAIDYSNKDATGKRLEPRRRLEIQKLRRWQARSR 111
ApeTFB1 GEDIIGTESDVKAYTPEERQQKTHYGGPLKYSHHYMGVEASLEH----PRDHGPKGIKQRKILPRRPPRLSAR 111
IhoTFB1 LDRYVDQGPEWRAFTPEERERRGRIGAPLSPTLHDHGLSTVIDHRDRDALGKRLSPRKRQEVQRLRKWQLRAR 110
DkaTFB1 EEKVIDERPEWRAFTSEERGRRARTGGPVTATVHDMGFATTIDYTDRDAAGRRLT-EKKHELVKLRKWQARTR 123
HbuTFB1 EDTVIDTGPEWRAYTPEEKTKRSRVGSPLTHTLPDYGVLTTISG-YRDATGRKLEARLRIEASRLRRLQAKLR 120
TtxTFB2 RENVVDLGPEWR------RPESSRAFVGTSGTELGDIKVGKPKLSDKLRAIKLREAVKPIPSPS--------- 101
PcaTFB2 KEGVADLGPEWR------KPEASRAYAGPIGSSIGDIERGNVKITDKLRAIMLGKFSKPMSTPL--------- 99
PaeTFB2 KDGIADLGPEWR------KPESSRAYSGPIGSSIGDIEFGNVKVADKLKAISLKKFSRPISTPL--------- 99
ParTFB2 REGIADLGPEWR------KPESSRAYSGPMGASIGDIEYGNVKITDKLKALSLKKFSRPISTPL--------- 99
PisTFB2 REGVADLGPEWR------KPESSRAYSGPIGSSMGDIEFGNVKIADKLKALSLKKFARPISTPL--------- 99
TneTFB2 RESIADLGPEWR------RPESSRAHAGPVGSSLGDIEYGNIKIMDKLKAISLKKFARPLTTPL--------- 99
TpeTFB2 PESLLFTGRSAR------SKEEFARKVHDYPVSRGIDTTPQEISEKYKRAYYESLGK---------------- 81
CmaTFB2 SEHEIDKGPEWR------GVDKNGKSLSRASPVSASSPGFNMYVNIVSVNRPKFRIIPSLSLFNSS------- 90
SmaTFB2 EDHLIDHGFEHR------YIDNFENARTSGSYTFRVHDSGIGSTEFKARYAGKWLNMSRLQKK--IRVEKR-- 99
DkaTFB2 FMRHLKQGRSWV--RVNSEIKVEKKDRKLKKALQELNEYLKLINPPPAVKETAAKLVHEAVK----------- 104
HbuTFB2 DDAPLDTGPEWR-----SYTEEDRLLRSRVGAPLTERVHDKGLTTYVQRNRRDPRAMKIVQLQTQLRTHG--- 104
SacTFB2 DEPLIDQGPEWR--AYTTEDKVERERTGSPLTAKVHDFGITTKIGYTKIKDRIKVHKLRLMQNKIRVSAR--- 97
StoTFB2 DEFVIDQGPEWR----AYTEQDRLERERTGSPITAKVHDFGITTTIGYSRSSNKTKIEKLRAIQNKLRVSPK- 100
SsoTFB2 EDSAVDPGPDWR--AYNAKDRNEKERVGSPSTPKVHDWGFHTIIGYGRAKDRLKTLKMQRMQNKIRVSPK--- 98
MseTFB2 EENVPDPGPEWR--EFDAGDRNKKRRVGAGITDRVHDKGFSTIIGGGRVKDKMKAIRLQRLQNKSRVTSK--- 97
TtxTFB3 ---------- 27
PaeTFB3 ---------- 27
ParTFB3a ---------- 39
ParTFB3b ---------- 33
ParTFB3c ---------- 43
PcaTFB3a ---------- 28
PcaTFB3b ---------- 30
PisTFB3 ---------- 27
SsoTFB3 ---------- 29
StoTFB3 ---------- 29
SacTFB3 ---------- 29
MseTFB3 ---------- 32
ApeTFB3 ---------- 34
HbuTFB3 ---------- 39
TtxTFB4 ----------GGAVASGAYATRRDDDHAPIFGSTLFDWHDIRGLVVPKYKVY 77


][cyclin - fold
BH1
BH2
BH3
hTFIIB* --------HHHHHHHHHHHHHH----HHHHHHHHHHHHHHHHH-------HHHHHHHHHHHHHHH------
TtxTFB1* -----HHHHHHHHHHHHHHHH----HHHHHHHHHHHHHHHHHHH------HHHHHHHHHHHHHH-----
yTFIIB DKK--DNEVQAAFAKITMLCDAAELPKIVKDCAKEAYKLCHDEKTLKGKSMESIMAASILIG--CRRAEVA 186
hTFIIB SSS--DRAMMNAFKEITTMADRINLPRNIVDRTNNLFKQVYEQKSLKGRANDAIASACLYIA--CRQEGVP 174
TtxTFB1 VQTSYERNFVQAAQELERLKSSMGVPRPCVEQALEIYRQALEKELVRGRSVEAMAAAALYMA--CRMLKTP 203
PaeTFB1 VQTSYERNFIQAAQELERLRSAMGIPRPCIEQALEIYRQALEKELVRGRSVEAMAAAALYMA--CRMMKMP 204
TpeTFB1 AQSSMERNLAQAVIELERLGAQLGLPKAVLDRALDIYKSALDSNLVRGRSIESVMAAAVYAA--CREMRLP 196
PisTFB1 VQTSYERNFIQAAQELERLRSSMGVPRPCVEQALEIYRQALEKELVRGRSVEAMAAAALYMA--CRMMKMP 204
PcaTFB1 VQTSYERNFIQAAQELERLKSSMGVPRPCVEQALEIYRQALEKELVRGRSVEAMAAAALYMA--CRMLKMP 202
ParTFB1 VQTSYERNFIQAAQELERLKSSMGVPRPCVEQALEIYRQALEKELVRGRSVEAMAAAALYMA--CRMMRMP 204
TneTFB1 VQTSYERNFIQAAQELERLRSSMGIPRPCIEQALEIYRQALEKELVRGRSVEAMAAAALYMA--CRMMKMP 204
CmaTFB1 VQTSYERNLVQATHELNRIAHQLGVPKSCMDEALAVYKQVLKSGLVKGRSVEAIIAACLHMA--CRMQGMP 207
SsoTFB1 IQSSIDRNLAQAMNELERIGNLLNLPKSVKDEAALIYRKAVEKGLVRGRSIESVVAAAIYAA--CRRMKLA 180
StoTFB1 IQSSIDRNLAQAMNELERIGNLLNLPKAVKDEAALIYRKAVEKGLVRGRSIESVVAASIYAA--CRRMKMA 179
SacTFB1 IQSSIDRNLAQAMNELERIGNLLNLPKAVKDEAALIYRKAVEKGLVRGRSIESVVAASIYAA--CRRMKMA 182
MseTFB1 IQSSIDRNLAQAMNELERIGNLLGLPKSVKDEGALIYRKAVEKGLVRGRSIESVVAAAIYAS--CRRMKIA 176
SmaTFB1 IQSSIERNLAQAMNELDRLSDQLNLPRNVKEEAARIYRRAVEKGLVRGRSIESVIAAAIYAA--CRELKIP 180
ApeTFB1 PLTSVDKNLQTALSLINEVASRMGMPEIVVEDASKIYREAMEKGLTRGRSIESIVAASLYAA--SRIHGLP 180
IhoTFB1 IQTGMDRNLTIAMNELDRMANLLNLPKQIKEEAAVIYRKAVEKGLVRGRSIESVVAAVIYAA--CRIHHQP 179
DkaTFB1 IITSVERNLAQAMNELDRLSDILNLPSYVKEEAARIYRNAVEKGIVRGRSIESIIAATIYLA--CREMKVP 192
HbuTFB1 ATTSIEKNIEQAAREITRLIEALNLPRSVIDTAMMIYRQAAEKGLVRGRSLESMAAAAVYAA--CRIRGIP 189
TtxTFB2 -----ERLEAEMREFYESLRHILGIPRALVDEAIALYRKAYDAGYK-ASRREGYAAALYFAV--KRHGLGA 164
PcaTFB2 -----ERLELDAREFFESARVRIGLPKVVVDEAVALYREVYKAGFR-APRIEGYAAVLYFVA--KRHGLAS 162
PaeTFB2 -----ERVESDVREFLESAKEKLSIPKNVIEETVMLYKKLYDAGYK-APRLESYAAVLYFVL--KKHGIAA 162
ParTFB2 -----ERVESEIREFLESAKEKLNIPKTVVEETVILYKKLYDAGYR-APRLESYAAVLYFTL--KRHGIPS 162
PisTFB2 -----ERIEVDIREFLEAAKQKLNIPKAVIEDTVMLYKKLYDAGYR-APRLEGYAAALYFTL--KKHGVAA 162
TneTFB2 -----ERVEIDIREFLDSAREKLNLHPSIVEETVQLYKRLYEAGYR-APRLEGYAAALLFTL--KSRGVAA 162
TpeTFB2 -----ERKSASYLREVESMAKSMNLPPHLIEEVKEYFTKVQKEGLLRGRNRKAVAAAIIYYV-CSKEKNVS 146
CmaTFB2 -----ERNVMMLRSIAKQVIANAGLPESILDEVVLNYRLLMKMN-YRGKIKETAVALVYIAC---RRRNLP 152
SmaTFB2 -----NKIVEKALRHLNNYIRILNPPKYVSETAGLILQKSVEGKNYKDKTLKLLAIASLYIAY--KVHGIP 163
DkaTFB2 -----DKN----------YKESTLRKIIIAAIYLSY--RINDTP 131
HbuTFB2 -----QKKMIKLLQDLNHFVAKLNLPPRVAETMAKLVKKLYLMGVI-KKNNEHVYLAAAAVIA-SRIEGHS 168
SacTFB2 -----ERKLVTYLSVLNSEASKLNLPEHVKETASILIRRLIEEGKA--KRVEMYALIAAVIYYSCQVNRIP 161
StoTFB2 -----DRKLVTYLSVLNNEAAKLNLPEYVKETASLLIRKIIDEGKA--KRIDIYTLIAAVLYYSCQVNKIP 164
SsoTFB2 -----DKKLVTLLSILNDESSKLELPEHVKETASLIIRKMVETGLT--KRIDQYTLIVAALYYSCQVNNIP 162
MseTFB2 -----DKKLVTYLSILNSEASKLGLPGYVKETAAGVIKKLIESGLA--RRIDTYALIAATLFYVSRLYKIP 161
TtxTFB3 ----------G 28
PaeTFB3 ----------G 28
ParTFB3a ----------G 40
ParTFB3b ----------G 34
ParTFB3c ----------G 44
PcaTFB3a ----------G 29
PcaTFB3b ----------G 31
PisTFB3 ----------G 28
SsoTFB3 ----------D 30
StoTFB3 ----------D 30
SacTFB3 ----------D 30
MseTFB3 ----------D 33
ApeTFB3 ----------A 35
HbuTFB3 ----------D 40
TtxTFB4 EMLN-RRVSSGVETEIDKLASLLSLPKACSSTAKWIVM-----RIRYVRDVELTAAAALFIS--CRIVKTY 140




```
hTFIIB*   HHHHHHHH----------HHHHHHHHHHHHHHHH------HHH-----HHH--------HHHHHHHH----
TtxTFB1*  HHHHHHHHHH-------HHHHHHHHHHHHHHH--------HHHHH-----H--------HHHHHHHHHH--
yTFIIB    RTFKEIQSLIHV-----KTKEFGKTLNIMKNILRGKSEDGFLKIDTDNMS--------GAQNLTYIPRFC  243
hTFIIB    RTFKEICAVSRI-----SKKEIGRCFKLILKAL------ETSVDLITTGD--------FMSRFCSNLCLP  225
TtxTFB1   RPLDELIRYTKA-----SRREVARCYRLLLR--------ELNVRVPISDP--------VLYISRVAEQLK  252
PaeTFB1   RPLDELVRYTKA-----SRREVARCYRLLLR--------ELNVKVPISDP--------ILYISRIAEQLK  253
TpeTFB1   RTLDEIALYTRA-----GRKDVARCYRLLLR--------EASLKVPLPNA--------ADFVPRIGSLLR  245
PisTFB1   RPLDELVRYTKA-----SRREVARCYRLLLR--------ELNVKVPISDP--------VLYISRIAEQLK  253
PcaTFB1   RPLDELVRYTKA-----SRREVARCYRLLLR--------ELNVKVPISDP--------VLYISRIAEQLK  251
ParTFB1   RPLDELVRYTKA-----SRREVARCYRLLLR--------ELNVKVPISDP--------VLYISRIAEQLK  253
TneTFB1   RPLDELVRYTKA-----SRREVARCYRLLLR--------ELNVKVPISDP--------TLYISRIAEQLK  253
CmaTFB1   RSLDEISQYTRA-----PRKEIARCFRLIAR--------ELRIRLPLSDP--------RQYVPKIVEQLK  256
SsoTFB1   RTLDEIAQYTKA-----NRKEVARCYRLLLR--------ELDVSVPVSDP--------KDYVTRIANLLG  229
StoTFB1   RTLDEIAQFTKA-----NRKEVARCYRLILR--------ELDINVPVSDP--------KDYVTRIGSLLG  228
SacTFB1   RTLDEIAQFTKA-----NRKEVARCYRLILR--------ELDIEVPVSDP--------KDYVTRIGTLLS  231
MseTFB1   RTLDEIAQYTKA-----NRKEVARCYRLLLR--------ELNVDVPVSDP--------KDYVTRIGSLLG  225
SmaTFB1   RTLDEIAKHTKS-----SRKDIARCYRLLLK--------ELDIKVQTSDP--------IDFIPRIAHALN  229
ApeTFB1   HSLTDIIKAMKGNVDAETRRDVARSYRLLVR--------DLNIKIPVRKP--------ENFVYTIISALG  234
IhoTFB1   RTLDEIAKKLEV-----NRKEVARCYRLITK--------ELKLKVPIADA--------MDHIPRIGEALK  228
DkaTFB1   RSLDEITRHTRV-----NRKEIARCYRLLLR--------ELHIKVTTTDP--------VDYVPRIVHGLG  241
HbuTFB1   RSIDDIAELVKG-----GRKEVARCYRLIVR--------ELRLRMPIVDP--------VRYVSRITSALR  238
TtxTFB2   VTYRSLIEKAGL-----DRSRFMSAYMEFLK-----YMSKLGEKVPRVDP--------RVYLPRIVSALK  216
PcaTFB2   VTLKALTEKLGI-----DRSALISAYMELMK-----VALSLGIRPPKADP--------KIYIPRIVASLG  214
PaeTFB2   VTLKKIIEELGI-----DRSNFISAYMELMS-----TANKLGIKPPRVDP--------RVYIPKIVSALG  214
ParTFB2   VTLKKIVEGLGI-----ERSSFISAYMELMK-----VASQLGIKPPRVDP--------KVYIPKIVSALG  214
PisTFB2   ITLRSLAENLGI-----NRSSFISAYMELMR-----VANKLGIKPPRIDP--------KIYIPRIVSALG  214
TneTFB2   VTPKTLYEQLGV-----DRSAFQSAYMELLK-----VANSLGLRPPRTDP--------KIYIPKIVAALG  214
TpeTFB2   VSLKDLEGVAGV-----SENDVKKAYRVLLK----KGVFELGQPHVSKPS--------KYVYKIVGSHEL  199
CmaTFB2   CTMKDLLRNSDV-----DIKGFNKAYMHIAN------LMNIKGIYNDEQL--------INMTLRIVNMIN  203
SmaTFB2   KSAKMFAKELGI-----TLKDLWRAEKKIHD-----NVKDINKMIKKDEP--------ENYVPYIVNRLS  215
DkaTFB2   KSLKVFISELNV-----DEKDLWEGVRLIRE-----SNSNIKISRENFDP--------RRYVGYIIRKLN  183
HbuTFB2   LTMRDVAETLGL-----SRQEIWKAYRKIVT-------KLKVRVATPPRP--------QMYVSKIASKLK  218
SacTFB2   KLLNEIKTLYSL-----SQADLWKALEKVQE------VAKSVKVKPNVTP--------IEYIPKITERLG  212
StoTFB2   KSLQEIKNNYGI-----SSSELWKALERVQK-----VAKSSLEFKPNIKP--------TEFIPKIVEKLN  216
SsoTFB2   RHLQEFKVRYSI-----SSSEFWSALKRVQY-----VANSIPGFRPKIKP--------AEYIPKILYKLN  214
MseTFB2   LQLNEIKKLFNI-----DSSAFWKASTRVQN------VIQKSTFMSNFGG--------PSPLEHIPDIVN  212
TtxTFB3   PVLYVPRIKY-------IQLGRNKDIASKILLM------QLNQKVIEKKYSYSEKILMYIKVLCRELDLP  85
PaeTFB3   PEVVPPRPRQ-------VLPVPVRHKLIMLALE--------QENKKSIKKKYSEMVKMYTGKVAEALGVP  83
ParTFB3a  PVLVPPVTKEA--------PRPPPRYRLIMTAL--------ERESRRSVRRYSEVVKMHLGKVAKALGAE  94
ParTFB3b  PVLMPPVLKEA--------PPPTPRYRLVVLAL-------EREGRRSVRRRYSEIVEMHLGKVAVALGAE  89
ParTFB3c  PEMLPPRLRQV--------PVPVHKGKIVLTIL-------EKEDKTTIKKRYSELVEHYVAKIAAALARK  99
PcaTFB3a  PVYMWPIRRVD---ERLAEKAAELSKGLRLALVKRGVPLREWLELEGARAWLEQKRRLTAWIPPSKRVEY  96
PcaTFB3b  YEALPPIVTKV--------QPIKKIKQILMIL--------EKEKREIIKTKYSDIVKYYIAKICRELGSP  85
PisTFB3   FEVVAPVSKQN--------MSVIRRRTIFLSL--------EKENKKTINMKYSDIVKFYINKISTELAIE  82
SsoTFB3   NIYYNGQNESE--------STEIISINNKFY---------KDDIQVKELRVKNFLKNRIETKKIDQYEII  83
StoTFB3   KIYYEDSTSYD--------ESVYFPETLYFDFF-------KTKRIKEFKNKTLKGRRNKLQSTIIYNGSV  85
SacTFB3   KIYDDHNYIDN--------EIMIKIQSTFTNVTILTY---KDKIEKIDKIIKFNSKLSKKQLPKSRKLLN  89
MseTFB3   RVYDYRPMFME--------EVLRSSQNRSLLYKNREIESFKEFMHLLKKSRFLRKYKGVKLNVSLNQGSI  95
ApeTFB3   EQLIDDSAPQH--------HSGEDHTPGPPLRL----------------SRLPRGVRRAYRRALRRGHVV  81
HbuTFB3   DLIVSYSPGPP--------LIVSHAGSLPVEKR--------------YTITEANVVRYEFAGLKLYEKLA  88
TtxTFB4   ADF----KFKGF-----DAEKLRRKIVLLQKLA------KLSMPPPPPVQ--------VIHHLVSEFKLG  187
```




```
hTFIIB*  HHHHHHHHH----------------HHHHHHH---------HHHHHHHHHHHHHH------HHHHHHHH--
TtxTFB1* --HHHHHH--------------HHHHHHHHHH---------HHHHHHHHHHHHHHHH-----HHHHHHHH--
yTFIIB   SHLGLPMQVTTS----------AEYTAKKCKEIKEIAGKSPITIAVVSIYLNILLFQIPITAAKVGQTLQV  304
hTFIIB   KQVQMAATH----------------IARKAVELDLVPGRSPISVAAAAIYMASQASAEKRTQKEIGDIAGV  280
TtxTFB1  LTGDVVKT--------------AIDIINRAKKAGLTAGKDPAGLAAAAVYIASLLHGDNRTQKDFAVAAGV  309
PaeTFB1  LSGEVVKT--------------AIDILQKAKKAGITAGKDPAGLAAAAVYIASLMHGDNRTQKDFAVAAGV  310
TpeTFB1  LSGATIKR--------------AMEIIDQARNAGLTAGKDPAGLAAAAIYIAALQNGEMRTQKEVARAAKV  302
PisTFB1  LSGEVVKT--------------AIEILQKAKKAGITAGKDPAGLAAAAVYIASLLHGDNRTQKDFAVAAGV  310
PcaTFB1  LSGEAIKL--------------AIDILQKAKKAGITAGKDPAGLAAAAVYIASLMLGDNKTQKDFAVAAGV  308
ParTFB1  LSGEVVKA--------------AIDILQRAKKAGITAGKDPAGLAAAAVYIASLMHGDNRTQKDFAVAAGV  310
TneTFB1  LSGEVVKT--------------AIDILQKAKKAGITAGKDPAGLAAAAVYIASLMHGDNRTQKDFAVAAGV  310
CmaTFB1  LPGDIAKE--------------AIRVLEEAKDKGLTAGKDPAGLAAAAVYIASLLKGEVRTQKEIAQAAQV  313
SsoTFB1  LSGAVMKT--------------AAEIIDKAKGSGLTAGKDPAGLAAAAIYIASLLHDERRTQKEIAQVAGV  286
StoTFB1  LSGSTMKM--------------AIDIIEKAKESGLTAGKDPAGLAAAAIYIAALLNDERRTQKEIAQIAGV  285
SacTFB1  LSGITMKH--------------AAEIIEKAKNSGLTAGKDPAGLAAAAIYIAALLNDERRTQKEIAQVAGV  288
MseTFB1  LSGASMKL--------------AAEILEKAKNVGLTAGKDPAGLAAAAIYIAALLNEERRTQKEIAQVAGV  282
SmaTFB1  LSGSVMKK--------------AAEILHRARSLGVTAGNDPAGLAAAAVYIAAQLSGERRTQKEIAHVAGV  286
ApeTFB1  LPEHVAIE--------------AIKIIDLSRKKGLTAGKDPGGLAGAAVYLAALKHGIRKTQKEIAHVVGV  291
IhoTFB1  LRGDIIEY--------------AMKIMEKIKGHPITAGKDPAGIAAAVIYIAVMQKGERRTQKEIANVAGV  285
DkaTFB1  LPGQAVKI--------------AIEIINKAKEQGVTGGKDPAGLAAAAVYIAAEQLGEKRTQKEIAHVAGV  298
HbuTFB1  LSPAVEKR--------------AAEILIRARKMGLTAGKDPAGLAAAAIYIAALELGERRTQKEIAAAAGV  295
TtxTFB2  LDGEIAAQVQKIA---------SDLLKYIMSSPRIRNGRKPQALAAVAVYYACFIAGLEVTQKDVARAAES  278
PcaTFB2  IGDEKSAEVQRVA---------VDILRYIISSPRIRNGRKPQVLAAAAVYYACFIAGVEVTQKELARAADS  276
PaeTFB2  IGDEKSAEVQRIA---------VDILRYIISVPRIRNGRKPQVLAAAAVFYACYIAGVEVTQKEVAKAADS  276
ParTFB2  LGDEKSAEVQRIA---------VDILRYITSIPRIKNGRKPQVLAAAAVFFACYIAGVEITQKEVAKAADS  276
PisTFB2  IGDEKSAEVQRIA---------VDILRYILSSPRIRNGRKPQVLAASAVYYACYIAGVEVTQKELAKATDT  276
TneTFB2  IGDEKTAEVLRIA---------VDILRYIMSSPRIRNGRKPQVLAAAAVYYACYLAGVEVTQKELAKAADS  276
TpeTFB2  QEVRSNIQL-------------KLLADFADSLGNLLQGKKPRGIAAAAVYIIANLLGYRKTQSLIAKIAGV  257
CmaTFB2  INSNIKHKLMLL----------IRDMIDKGKATSLFNGKTFTSTIAAMVYLSLLIYNVKIRQREIASMAGV  264
SmaTFB2  LSERVQYL--------------ANYIIQLSKKAGLNNGKSSVGLATAAVYIASILLNEKRTQIDVAKTVNV  272
DkaTFB2  APPIVET---------------MANYIISVTEEQGFSGKSPTSLAAAAVYLAGILTNNKRNQVEVAETVGL  239
HbuTFB2  LSGEVEAL--------------ATRFTIMLTKTGIAQGKPPEALAAAAVYVASILLDEKRNQQRVAQAIGV  275
SacTFB2  LPAYVSTK--------------ASELVDIMYKNGLTSGKGYTALAAASVYLISTLMDVKKTQKEIADSLSI  269
StoTFB2  LPPYISTK--------------ASELVDLMHKNGLTSGKGYTALAAASVYLVSTLMDAKKTQKDIANALNI  273
SsoTFB2  LPPIIGTK--------------ASELVDLMHKQGLTSGKGYLSLSAASVYLISALMDIKKTQKEVADSLDI  271
MseTFB2  RAKLPPHVETM----------AAEIASILIKNAITSGKGHLSVAAAAVYLASTILDHKKTQKELAEVLNI  272
TtxTFB3  EEVFRSSIQ----------------VLKDINK-VKIQGKNPKVIAATIVYLISNKFKLNINKEQIARRLGI  139
PaeTFB3  E-VTVVALQ----------------IFQRLDK-RIYQGKSPRVIAAAVAYLAAERLGIYIHKQVIAKILGV  136
ParTFB3a V--AAVALD----------------IFKRLDK-RVYQGRSPRVVAATLAYLAAERLGIYVHKRVIAEILKV  146
ParTFB3b V--AAVALE----------------MFRRLDK-RVYT-RSPRVVAATLAYLAAERLKLYVHKQTIAEILKV  140
ParTFB3c D-LEMVALE----------------IFRNLDK-RVYQGKTPKAIAAALVYLAAERAGIHIYKQTIARILGI  152
PcaTFB3a Y-IEAAAAKLGLGRAAL---EEALALYRRLDR-RALV-KSPRVVAAVIYTATCTPARAAAEALGVSPISVK  161
PcaTFB3b E-LEREALE----------------LLSTIDK-RVWQGKSPRVIAASLVYLAAERLNYHFHKRQIGEIVKV  138
PisTFB3  N-IREVAWD----------------LFQKLDK-RVYQGKNPRVIAASVVYLAAEAMGVQIYKQTIAKIVNI  135
SsoTFB3  LRSMLLDSQYKK------------IYKLLYDE-GILSGLKAKSKLGLLIYFRFALNNGYLHLLEKFDIKNE  141
StoTFB3  IKETSLNAMKFLENNEK--------LLLLYDTIDNLPVFHSKSVKYKLAIALYFYDKKEFNRLSRYLNISN  148
SacTFB3  YNNAIIRSSSATIIKYLDFNEKLLLVYDIIDTIPILNNISIKYKVALAMYFYDKKTFNKIMNNLEISNKYF  160
MseTFB3  QGSKIYSQISLDALYKIESDKVAIKIYKYLEKMGIFSGLKFKTRVLLTYYLVYGNDKIRIKRILKTYYSSE  166
ApeTFB3  VGEGSMPYSDLRAV-EMSSDESIAAILGMVDSIPGLARRRPRVKVGVAFYISYRLRGYSKSSALRTAARSS  151
HbuTFB3  KLLGSRSVERVREAAAGDRHALSVIASNKCLEDVMRSLPEPERGVVVELALSFMRGEYPLLSIVAERYGVS  159
TtxTFB4  PEVVRDSLD----------------ILEVL---SPILGRGKIAQLVAFI-LAARARGHKIDVRAVADASWS  238
```


```
hTFIIB*   -HHHHHHHHHHHHHHHHHH----------------------
TtxTFB1*  -HHHHHHHHHHHHHHHH------------------------
yTFIIB    TEGTIKSGYKILYEHRDKLVDPQLIANGVVSLDNLPGVEKK  345
hTFIIB    ADVTIRQSYRLIYPRAPDLFPTDFKFDTPVDKLPQL-----  316
TtxTFB1   TEVTVRNRYKELAKALDIKIPVK------------------  332
PaeTFB1   TEVTVRNRYKELAKALNIKVPVK------------------  333
TpeTFB1   TEVTVRNRYKELVKKLDIKLPVSKK----------------  327
PisTFB1   TEVTVRNRYKELAKTLNIKVPVK------------------  333
PcaTFB1   TEVTVRNRYKELAKALNIKVPIK------------------  331
ParTFB1   TEVTVRNRYKELAKALNIKVPVK------------------  333
TneTFB1   TEVTVRNRYKELAKALNIKVPVK------------------  333
CmaTFB1   TEVTVRNRYKELAKELNIKIPIK------------------  336
SsoTFB1   TEVTVRNRYKELTQELKISIPTQ------------------  309
StoTFB1   TEVTVRNRYKELTQELKIQIPNQ------------------  308
SacTFB1   TEVTVRNRYKELTQELKIQIPSQ------------------  311
MseTFB1   TEVTVRNRYKELIQELKIEIQNQ------------------  305
SmaTFB1   TEVTVRNRYKELAKELGIDLPTT------------------  309
ApeTFB1   TEVTIRNRYKEIAQALGIEEELEEKGGEEKG----------  322
IhoTFB1   TEVTVRNRYKEIMKVLNEMDLEEIEKEVSKK----------  316
DkaTFB1   TEVTVRNRYKELAKILGLEYE--------------------  319
HbuTFB1   TEVTVRNRYKELVQKLNITLPAQ------------------  318
TtxTFB2   TETPIRELLNELGKRLYIEISI--------------------  300
PcaTFB2   TEGPIRELLRELADKLYIELTV--------------------  298
PaeTFB2   TEGPVRELLKDLSDVIYVEVTV--------------------  298
ParTFB2   TEGPVRELLKDLSNVLYIEITV--------------------  298
PisTFB2   TEGPIRDLLKELSEILYIEVTV--------------------  298
TneTFB2   TEGPVRELLKELSERLYIEITL--------------------  298
TpeTFB2   STLTLRRVVKEIDENLDIYIEI--------------------  279
CmaTFB2   TDVTIRNRYNELLKRLNFRVII--------------------  286
SmaTFB2   TDVTIRNRYSDIVRSFDITISI--------------------  294
DkaTFB2   SDVAIRNAYGSIIKDVDIEVLL--------------------  261
HbuTFB2   TDATIRNRYRDIVDNFYIEVRL--------------------  297
SacTFB2   TEVTIRNRYREIIKAFDIEVKL--------------------  291
StoTFB2   TEVTIRNRYKEIISNFEIEVKL--------------------  295
SsoTFB2   TEVTIRNRYKDIVDNFDIVVTL--------------------  293
MseTFB2   TEVTIRNRYKEIVNSLDVEVKL--------------------  294
TtxTFB3   SKLTIRDTATQLRKLCKELNY--------------------  160
PaeTFB3   SKFSIRDTASRLRKYVPNIKETT------------------  159
ParTFB3a  SKFTIKDTVWRLRRYLQEA----------------------  165
ParTFB3b  SKFTIKDTAWRLRRHLQEA----------------------  159
ParTFB3c  SKFTIRDTVTRLRGHVASINTSS------------------  175
PcaTFB3a  ATVRKLKLAVKICVRRGVPRGA------------------  183
PcaTFB3b  SKFSIRDTVVKLRKYV--------------------------  154
PisTFB3   SKFTIRDTVSKLRRHVSTTH----------------------  155
SsoTFB3   TLKKILKRIGRKRLTLLFDKLNEESDRI-------------  169
StoTFB3   KYMNKILSKIKIKEKTKIQHILKNKIGKRTNNSFYNI----  185
SacTFB3   NKILRRLNSKEKMIIMEKVINLLEQRVPSQTLKTM------  195
MseTFB3   ENIKRIIKQIPFNVRLEIMRMVEESKIKN------------  195
ApeTFB3   GASVAALERVEKAYRREIENLIMAASMGGERV---------  183
HbuTFB3   RARVRSLARRVRKCMGQPSLSEALASMSIAASLTG------  194
TtxTFB4   QPDLIVDALKTVARAVK-----------------------  255
```

FIGURE 3. Multiple sequence alignment of TFB1-4 homologues from Crenarchaeota.
The alignment was performed using ClustalX. Sequences from various crenarchaeal species are presented: Pae, *P. aerophilum*; Pis, *P. islandicum*; Pca, *P. calidofontis*; Par, *P. arsenaticum*; Tne, *T. neutrophilum*; Ttx, *T. tenax*; Cma, *C. maquilingensis*; Iho, *I. hospitalis*; Tpe, *T. pendens*; Sso, *S. solfataricus*; Mse, *M. sedula*; Sac, *S. acidocaldarius*; Sto, *S. tokodaii*; Ape, *A. pernix*; Sma, *S. marinus*; Hbu, *H. butylicus*; Dka, *D. kamchatkensis*, yTFIIB: yeast TFIIB, hTFIIB: human TFIIB. For comparison, sequences of TFBs/TFIIBs from human and yeast are shown alongside structural data. The domains of TFB/TFIIB are depicted according to BUSHNELL et al. (2004). Domain boundaries are depicted by brackets; Zn-ribbon, B-finger, linker and two cyclin (HTH) domains. Secondary structure of Zn-ribbon and cyclin-fold is shown according to crystal structure of N-terminal domain of *Pyrococcus furiosus* (ZHU et al., 1996) and C-terminal domain of *Pyrococcus woesei* (KOSA et al., 1997) with β-sheets and α-helices despicted as red arrows and green bars, respectively. TtxTFB1*: predicted secondary structure of Zn-ribbon [(E): β-sheets] and cyclin-fold [(H): α-helices] of TFB1 from *T. tenax* by the program HHpred (SÖDING, 1995) based on the determined secondary structure of the Zn-ribbon (1dl6_A, Probab = 99,65, E-value = 4,4e-17, Indentities = 33%) and cyclin-fold (1c9b_A, Probab = 100, E-value = 0, Indentities = 31%) of human TFIIB (hTFIIB*). The Cys residues that ligand the zinc atom in the Zn-ribbon motif are highlighted in yellow and the conserved residues of the B-finger motif is highlighted in black. The analogous residues to E46 located in the B-finger of *Sac*-TFB1, important for the interaction with RNAP (BELL & JACKSON, 2000) is indicated by (+) (In *T. tenax*: E72); the analogous residues to L38 from *Sso*-TFB1, important for interact with RpoK (MAGILL et al., 2001) is indicated by dark blue (In *T. tenax*: I62). The two pairs of amino acids (G-K(R)) involved in the interaction with TBP located between helix I and II of both cyclin domains (BALIGA et al., 2000) are indicated by pink. The three amino acids located at the second cyclin fold domain which make base-specific contacts with BRE (MICORESCU et al., 2008) are indicated by light blue (In *T. tenax*: Q300, V312 and R315).

The B-finger is a conserved motif present in all crenarchaeal TFB1 and TFB2 homologues with only few exceptions (TFB1 *A. pernix*, TFB2 *T. pendens* and TFB2 *D. kamchatkensis*). The first cyclin domain is present in all crenarchaeal TFB1. Most of the TFB2 homologues from the Thermoproteaceae lack the first cyclin domain, except *T. pendens*. In all other crenarchaeal TFB2 with the only exception of *M. sedula* the first cyclin domain is present (Figure 3 and 4). The second cyclin domain is absent in TFB2 of *H. butylicus* and most Thermoproteaceae, except in *T. tenax*, *P. islandicus* and *T. pendens*.

Interestingly, TFB3 possesses half size of TFB1 and 2 and the B-finger motif, the linker and the first cyclin domain is missing in all TFB3 homologues (Figure 3). TFB4 contains no B-finger motif even though the protein has similar length as TFB1 and TFB2 and also no cyclin domains can be identified. However, the secondary structure prediction reveals two times five α-helices in TFB4.

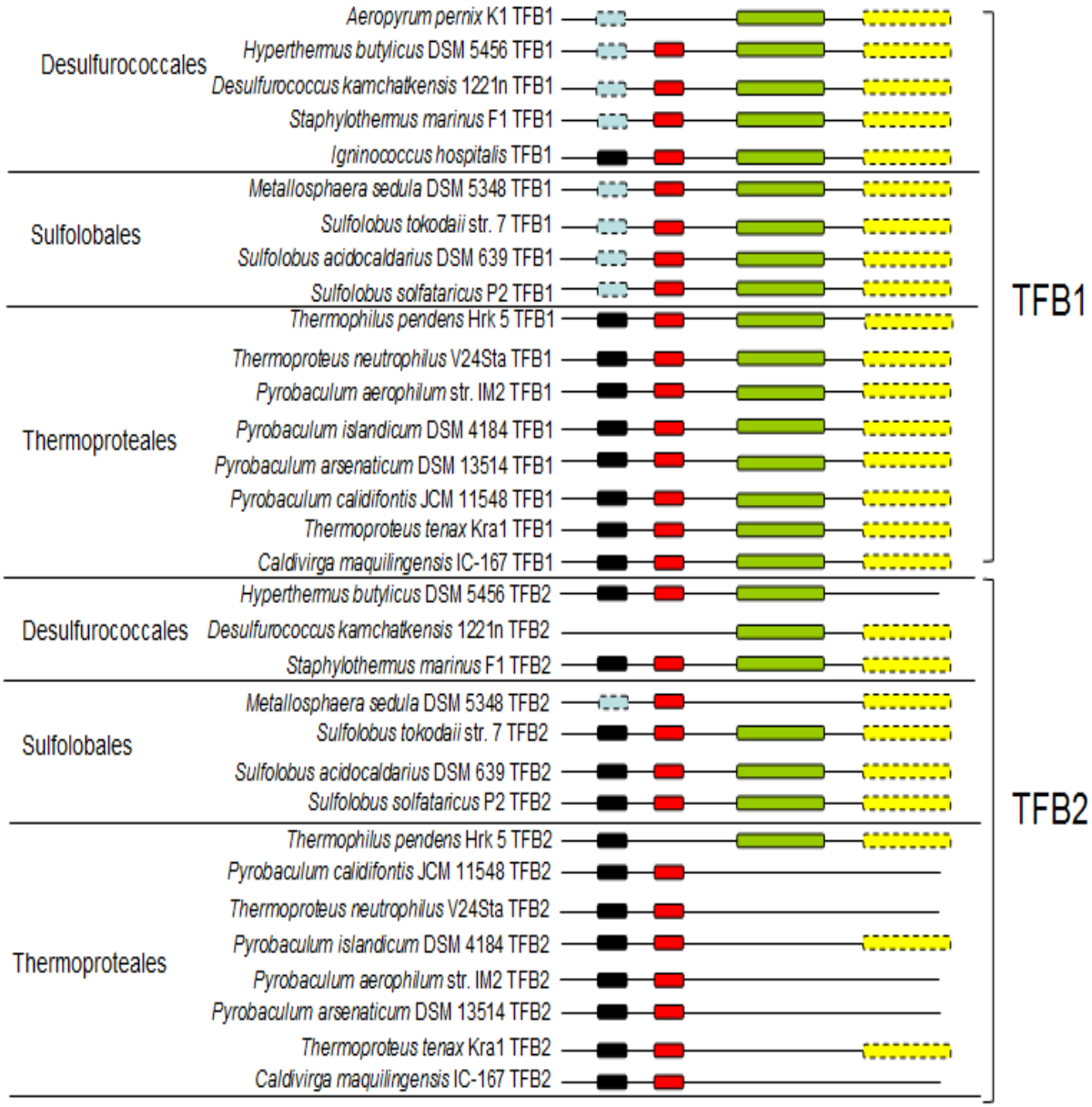


FIGURE 4. Domain and structural analysis of TFB1 and TFB2 homologues in Crenarchaeota. The CD-search tool at NCBI was used for the analysis: canonical Zn-finger: black square; non-canonical Zn-finger: blue square with dashed lines; B-finger: red square; first cyclin domain: green square; second cyclin domain: yellow square with dashed line.

In order to gain more insights into a putative functional association of TFBs, the genome context was analyzed. The *tfb1* gene is in close neighbourhood to the gene *rfc* coding for replication factor C in most Thermoproteaceae, except *P. arsenaticum* and *C. maquilinguensis*; to gene *rpoP* encoding subunit P of DNA-directed RNAP in Sulfolobaceae and Desulfurococcaceae or gene encoding lysil-tRNA synthase (COG1384) in *H. butylicus*. In addition, a hydrolase (COG0537) is located downstream to *tfb1* in Thermoproteaceae except in *P. islandicum*, *T. tenax*, *T. pendens*, and *P. arsenaticum* (Figure 5A).

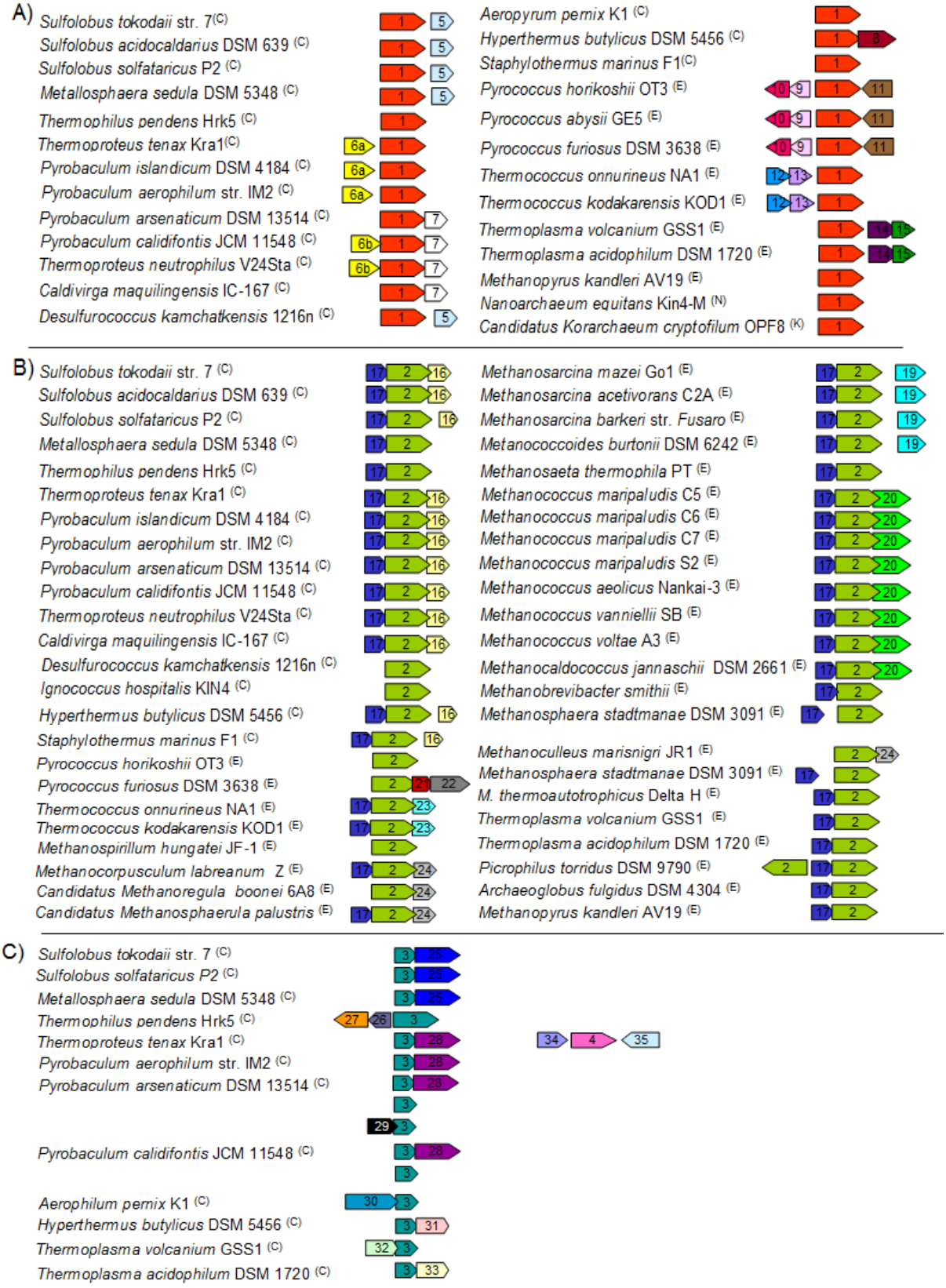


FIGURE 5. Genome context analysis of TFB1- 4 in archaeal genomes.
1: TFB1 homologues; 2: TFB2 homologues; 3: TFB3 homologues; 4: TFB4 homologues; 5: DNA-directed RNApol, subunit P (rpoP), COG1369; 6a and 6b: replication factor C, small subunit, COG0470 and COG2812, respectively; 7: hydrolase, COG0537; 8: lysil-tRNA synthase, COG1384; 9: ribonulcease P protein subunit, COG1369; 10: cell division protein, COG1537; 11: tRNA-Ser;; COG1996; 12: hypothetical DNA binding protein, COG2118; 13: hypothetical protein; 14:

Hypothetical protein, COG0518; 15: hypothetical protein, Q8U2Z2; 16: C34 subunit homologue, COG1846; 17: Gar1, COG3277; 18: tRNA, COG2419; 19: LysE type translocator; 20: type II secretion system protein E, COG4962; 21: hypothetical protein; Q8U2Z2; 22: hypothetical protein, COG2810; 23: metallophosphoesterase, COG0622; 24: tRNA-Ala; 25: ths, thermosome subunit, COG0459; 26: small nuclear ribonucleoprotein, LSM family, COG1958; 27: Met-adenosyltransferase, COG1812; 28: type II secretion protein E, COG0630; 29: hypothetical protein; 30: acylamino-acid releasing enzyme, COG1506; 31: predicted GTPase, COG1084; 32: molybdopterin synthase, COG1977; 33: hypothetical protein; 34: alcohol deshydrogenase; 35: aspartate aminotransferase.

In Euryarchaeota as well as in Crenarchaeota the gene *gar1* encoding the ribonucleoprotein Gar1, involved in rRNA modification, is located upstream of the *tfb2* gene and a gene coding for a homologue of the eukaryotic C34 subunit of the RNAP III is found adjacent to it in Crenarchaeota, and in Euryarchaeota a gene encoding tRNA-Ala in Methanomicrobiaceae; LysE type translocator in Methanosarcinaceae or type II secretion system protein E in Methanococcaceae. In *Picrophilus torridus*, the *gar1* gene is inserted between two *tfb2* genes, which arose from a gene duplication and are organized in a diverse organization (Figure 5B).

In summary, gene context analyses in Archaea illustrate a conserved organization of *tfb1* and *tfb3* genes only in the family of Thermoproteaceae (Figure 5A and 5C) contrasting with the high sequence conservation of TFB1 homologues within the domain of Archaea (Figure 3). In contrast, the genome organization of genes encoding *tfb2* homologues is conserved within archaeal domain (Figure 5B) but protein similarity is less conserved (Figure 3). The *tfb4* gene is embedded between two genes encoding alcohol dehydrogenase and aspartate aminotransferase, enzymes involved in the metabolism and non-related to informational processing mechanisms (Figure 5C).

3.2 HETEROLOGOUS EXPRESSION OF *T. tenax* GTFs IN *E. coli.*

In this work the *tfb2*, *tfb3* and *tfb4* genes were cloned into pET expression system via PCR mutagenesis. The primer sets used to amplify each gene are given in Table 1. In general, the genes were amplified by PCR (see above 2.4.6) and the PCR fragments were purified using the Promega kit (Mannheim, GER). Afterwards the PCR fragments and cloning vector (*pBlueScript* II KS(+)) were restricted using the respective restriction enzymes and later on purified from agarose gel and the ligation was accomplished overnight at 4°C. Transformation was accomplished using competent DH5α cells and the transformants were plated on LB agar plates supplemented with ampicillin. After incubation of LB-agar plates at 37°C overnight, colonies were screened for positive clones carrying the recombinant plasmid DNA by boiling PCR and restriction analysis of the isolated plasmid DNA. The DNA sequences of both strands of the gene of interest were confirmed by sequencing (AGOWA). Later on the insert

was excised from the cloning vector *pBlueScript* II KS(+) and cloned into the respective expression vector (Table 1).

The expression strain *E. coli* Rosetta(DE3) was transformed with the recombinant expression plasmid encompassing *tfb2*, *tfb3* and *tfb4* and the cells were grown overnight at 37°C (200 rpm). 2% (v/v) of the preculture was used for inoculation of prewarmed LB medium and incubated at 37°C in a rotary shaker (200 rpm). Protein expression was induced at OD_{600} = 0.6 – 0.8 by the addition of 1 mM IPTG and incubation continued for 3 – 4 hours (see above 2.3). In this work, 6xHis-*Ttx*-TFB2 and 6xHis-*Ttx*-TFB3 were successfully expressed as 6xHis-tag fusion proteins.

The recombinant proteins, 6xHis-*Ttx*-TFB2 and 6xHis-*Ttx*-TFB3 exhibited a good expression in the soluble fraction. The determined molecular masses (SDS-PAGE) approximately correspond to the theoretical molecular weight for 6xHis-*Ttx*-TFB2 (35 KDa, 313 aa) and 6xHis-*Ttx*-TFB3 (21 KDa, 180 aa) (Figure 6 and 7). 6xHis-*Ttx*-TFB3 was expressed also in the membrane fraction as shown in Figure 7, however sufficient enrichment in soluble fraction was observed from SDS-PAGE and the amount of soluble protein was enough for further purification steps (Figure 7).

Despite a lot of effort using different expression vectors (i.e. pET15b and pET24a) and expression strains (i.e. BL21(DE3) and Rosetta(DE3)), recombinant expression of 6xHis-*Ttx*-TFB4 was not successful. Also expression in the vector pQE-30, using the expression strain M-15 was unsuccessful as confirmed by western blotting using 6xHis-tag monoclonal antibody revealing no signal.

In order to achieve recombinant expression of *Ttx*-TFB4, modifications in the translational initiation and termination regions of the *tfb4* gene were introduced via PCR mutagenesis. The primer sets used for this approach are given in Table 1. The start codon GUG was changed to AUG, UAAUGA was introduced as translation termination sequence, and the third nucleotide, G, in the fourth codon and in the eigth codon was replaced by A, as recommended by MAKRIDES, 1996. The modified *tfb4* gene was cloned into different expression vectors (pET11c, pET15b, and pET24a) and expression was tested in different strains (BL21(DE3) and Rosetta(DE3)). However, all attempts to achieve expression were not successful.

3.3 PROTEIN ENRICHMENT AND PURIFICATION.

For further biochemical studies, 6xHis-*Ttx*-TFB2 and 6xHis-*Ttx*-TFB3 were purified to apparent homogeneity from crude extracts. Briefly, cells (10 g wet weight) were suspended in 20 ml of buffer and passed three times through a French pressure cell. Cell debris and

unbroken cells were removed by ultracentrifugation in order to gain a cell-free crude extract. The crude extracts were subjected to heat precipitation, at different temperatures for 15 min; 70°C for 6xHis-*Ttx*-TFB2 and 60°C for 6xHis-*Ttx*-TFB3. Further purification steps of the recombinant proteins 6xHis-*Ttx*-TFB2 and 6xHis-*Ttx*-TFB3 were achieved using affinity chromatography as described in 2.5.1.4 and 2.5.1.5, respectively. Samples containing the enriched protein fractions were analyzed by SDS-PAGE, pooled and used for further studies to unravel their biological activity (Figure 6 and 7).

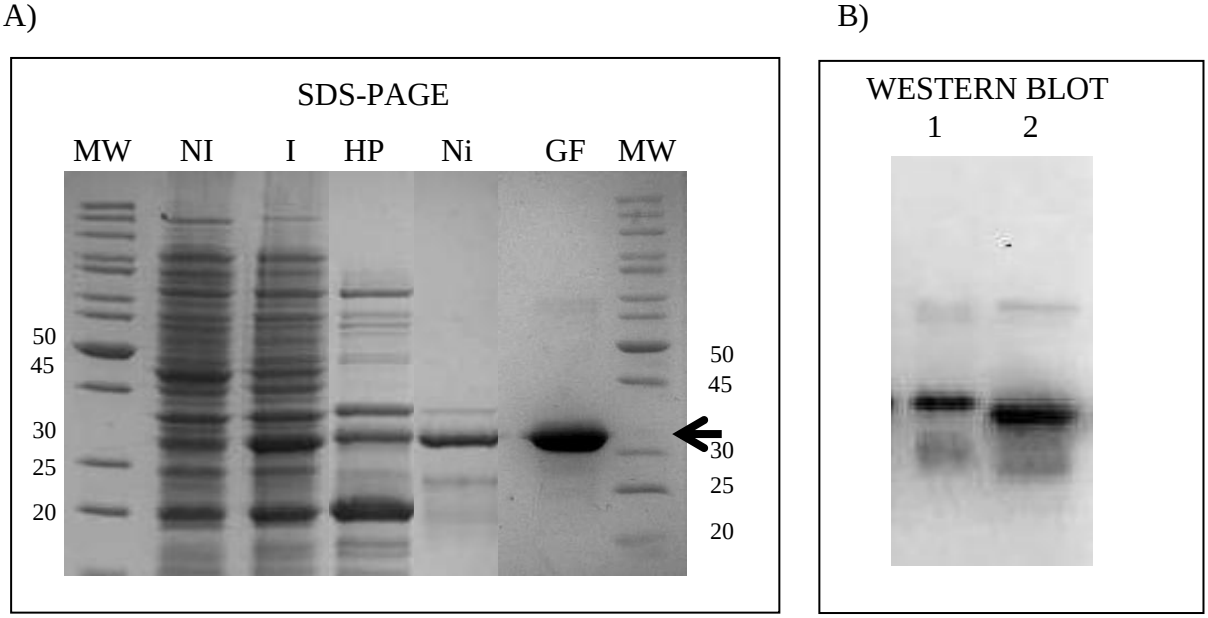


FIGURE 6. SDS-PAGE and western blot analysis of the purified recombinant 6xHis-*Ttx*-TFB2.
(A) Samples before induction (NI), after induction with IPTG (I), after heat precipitation (HP), after Ni-NTA column (Ni) and gel filtration were separated alongside the molecular weight marker (MW). The arrow indicates the position of 6xHis-*Ttx*-TFB2 and its determined molecular weight is 35 kDa. (B) Detection of the recombinant proteins by western blot. 6xHis-*Ttx*-TFB2 (lane 1) and *Ttx*-TFB2 (lane 2) were detected using *Ttx*-TFB2 polyclonal antibody (1:500) generated in this study. Mouse Rabbit-IgG monoclonal antibody conjugated to alkaline phosphatase was used as secondary antibody (1:10,000).

It is northeworthy that 6xHis-*Ttx*-TFB3 was unstable during purification as observed by the formation of degradation products after SDS-PAGE. Therefore, all the purification steps of 6xHis-*Ttx*-TFB3 required the addition of 20% glycerol in order to improve the stability of the recombinant protein.

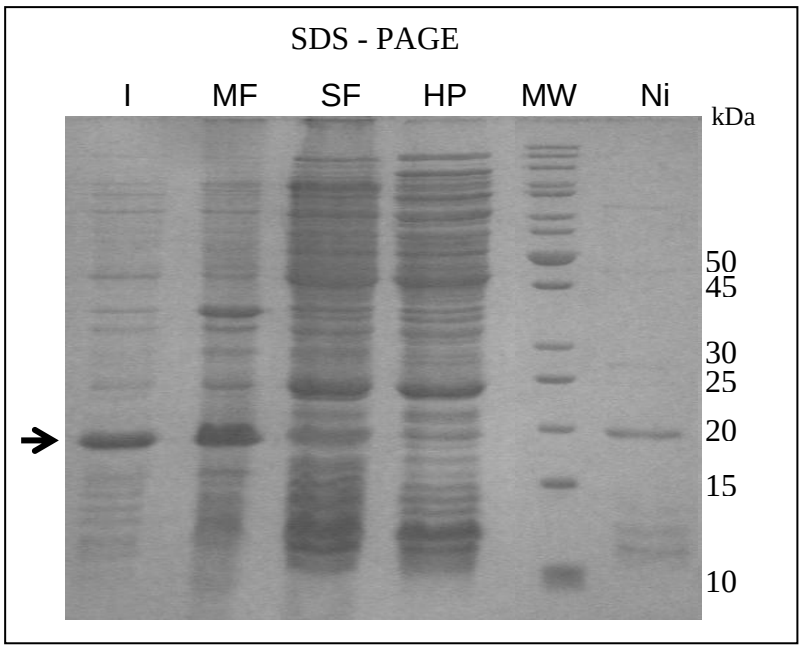


FIGURE 7. SDS-PAGE of the purified recombinant 6xHis-*Ttx*-TFB3.
Samples after induction with IPTG (I), membrane fraction (MF), soluble fraction (SF), after heat precipitation (60°C, 15 min) (HP) and after Ni-NTA column (Ni) were separated alongside the molecular weight marker (MW). The arrow indicates the position of 6xHis-*Ttx*-TFB3 and its determined molecular weight is 20,59 kDa.

The *tfb1*, *tfb2* and *tbp* genes have been cloned into the pET expression system via PCR mutagenesis, expressed and purified previously by F. BLOMBACH (2005). *Ttx*-TFB2 and *Ttx*-TBP were expressed in the soluble fraction and *Ttx*-TFB1 was expressed in inclusion bodies. The purification protocols were established by F. BLOMBACH. After purification by F. BLOMBACH the proteins (*Ttx*-TFB1, *Ttx*-TFB2 and *Ttx*-TBP) were stored in the presence of 10% (v/v) of glycerol at -80°C. The available proteins were used in this study for EMSAs and footprinting analysis.

1 mg of the purified proteins, *Ttx*-TBP purified according to the previously established purification method (BLOMBACH, 2005) and 6xHis-*Ttx*-TFB2 were used as antigen for the production of polyclonal antibodies (Eurogentec). As shown in Figure 6B and 8 6xHis-*Ttx*-TFB2, *Ttx*-TFB2 and *Ttx*-TBP were specifically detected by western blotting using the respective specific polyclonal antibody.

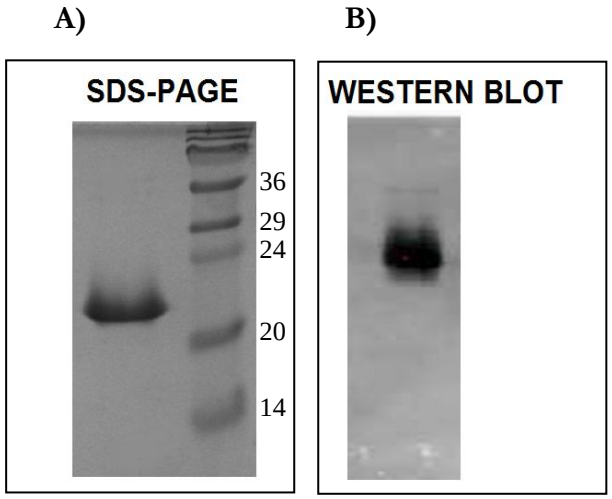


FIGURE 8. SDS-PAGE and western blot analysis of the purified recombinant *Ttx*-TBP.
(A) Recombinant protein was purified by heat precipitation (90°C, 15 min) and anion exchange chromatography previously by F. BLOMBACH, 2005. The purified protein was separated (12 % SDS-PAGE) alongside the molecular weight marker (MW). B) Detection of the recombinant *Ttx*-TBP by western blot. *Ttx*-TBP was detected using *Ttx*-TBP polyclonal antibody (1:500) generated in this work (2.5.2.3). Mouse Rabbit-IgG monoclonal antibody conjugated to alkaline phosphatase was used as secondary antibody (1:10,000).

3.4 DNA BINDING OF *T. tenax* GTFs.

3.4.1 DNA binding of *Ttx*-TFB1 and *Ttx*-TFB2 using a large-length *fba-pfp* promoter fragment.

DNA binding of *Ttx*-TFBs was tested using electrophoretic mobility shift assays (EMSAs). Large 265 bp DNA fragments comprising the transcription control regions of the *fba-pfp* operon were used as DNA probes in EMSAs in this work. The *fba-pfp* operon encodes enzymes involved in the central carbohydrate metabolism (CCM); the *fba* and *pfp* genes encode fructose-1,6-bisphosphate aldolase and PP_i-dependent phosphofructokinase, respectively, which are up-regulated under heterotrophic growth conditions (SIEBERS et al., 2004).

As shown in Figure 9, *Ttx*-TFB1 binds the *fba-pfp* DNA promoter fragment (12.5 fmol), however, binding appeared to be unstable during electrophoresis. No complex formation is observed in binding assays in presence of 30 ng *Ttx*-TBP alone and 12.5 fmol of the *fba-pfp* DNA promoter fragment, probably because the DNA-protein complexes are unstable during electrophoresis (Figure 9, lane 2).

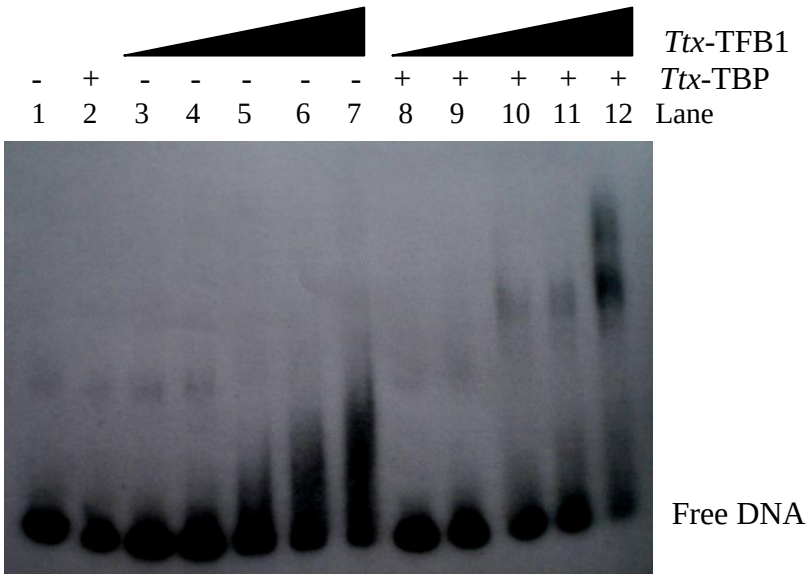


FIGURE 9. DNA binding of *Ttx*-TFB1 in the presence and absence of *Ttx*-TBP.
The DNA probe comprises the transcriptional control region of the *fba-pfp* operon (528 bp, Figure 10). Equal amounts (30 ng) of *Ttx*-TBP (+) and increasing amounts (0, 2, 6, 12, 20 and 30 ng) of the *Ttx*-TFB1 were added as indicated. Similar results were obtained using the promoter fragment of the genes encoding *tfb1*, *pps*, and operon *orf1155-lrp* as DNA probes.

Similar results were also observed previously by S. KONING, in the laboratory of Prof. Dr. BETTINA SIEBERS when DNA promoter fragments of different genes or operons *tfb1*, *gar1-tfb2*, *pps* and *orf1155-lrp* were used as DNA probes in EMSAs. The *tfb1* gene and *gar1-tfb2* operon encode *Ttx*-TFB1 and, Gar1 and *Ttx*-TFB2; respectively (Figure 5). The *pps* gene encodes the gluconeogenic CCM enzyme-phosphoenolpyruvate synthase, which is up-regulated under autotrophic growth conditions (SIEBERS et al., 2004; TJADEN et al., 2006). The operon, *orf1155-lrp* encodes for a protein of unknown function (open reading frame 1155, ORF1155) and a homologue of the leucine-responsive regulatory protein (Lrp). The DNA sequences of *fba-pfp*, *pps*, *orf1155-lrp*, *gar1-tfb2*, and *tfb1* probes used for EMSAs are shown in Figure 10.

At low amounts of *Ttx*-TFB1 (2 ng – 6 ng) the addition of *Ttx*-TBP (30 ng) shows no effect on DNA binding (Figure 10, lane 8 and 9). Notably, at higher amounts of *Ttx*-TFB1 (12 – 30 ng) the complex in presence of *Ttx*-TBP (30 ng) appeared to be more stable indicating that the presence of *Ttx*-TFB1 and *Ttx*-TBP stabilizes DNA binding. Similar results for DNA binding of *Ttx*-TFB1 in presence and absence of *Ttx*-TBP were also confirmed by S. KONING for the promoter fragments *tfb1*, *gar1-tfb2*, *pps* and *orf1155-lrp*.

```
>fba-pfp           CCGCCTCAATGGCGGCGTCGGTCGACTCGATATAACGATTTAACGCGCATCC
>pps               ------------------------------------------------------G
>orf1155-lrp       ------------------------------------------------------
>gar1-tfb2         ------------------------------------------------------
>tfb1              ------------------------------------------------------

>fba-pfp      AAGAGATATACTTCAATAAAATCATAAAACTCATATCCCATTATATTAATTGCTCTACTTAATAATAT
>pps          ATCCTGAAACTTTAGAGTAATTATATATTACGTGTTTTCTAGGCTATAGCCGGATCATCTATCTCTTC
>orf1155-lrp  ----CGACACCTTGGGCAGTTTGACCGAGGATGCATCTCAGTATATACAACCTCCGAGTGGAAAAGTC
>gar1-tfb2    --------------------GGGCGTACGTCGTGGATCCGCGGAGGATGGATAACACAGAAGATATCA
>tfb1         ------------------------------------------------------AATAAACTAAATAAC

                      BRE/TATA
>fba-pfp      ACTTTAGACAAAAAGATATTAAAATGGATAATTGCTCAAGGATCAatgGCAAACCTCACCGAGAATTCT
>pps          GAGTTTAATAGGTAGTATTAATACCACTCGTAATAACGCCCCatgGGGCTCATTCTATGGCTAGAGGAG
>orf1155-lrp  ACAGATTGAACATATAATTTAATAACCCAGATTATTGCTCGTCCgtgGACGAGATAGACAGGAAGCTTA
>gar1-tfb2    CAATTAGGGGGGAACCTTTAATAAACGCGCGGCTATAGACACTatgCTCAGACGTATTGGTATTGCTCA
>tfb1         TATATATGCGAAAAATTTATATACTTATAAAGTATTTATAAGGGatgTCGGCTCAGGGCA---------

>fba-pfp      TAAGGATATTCGCGAGGAGGGGGAAGTCCA---------------------------------------
>pps          ATCACGAAGAAAGATCACGCTCTAGTGGGCGGCAAGGGGGCAAACTTGGGCGAGGTCTCCAGACTCGT
>orf1155-lrp  TAGAACTTCTACAGATGGACGGCAAAAAGACGCTCCAGGAGCTGGCCGAGGCCGTGAATAGGCCGAAG
>gar1-tfb2    CCATTATACTAATCTAAAGAATCTGGTTGTAAAATTATCTACAGTCCCCCCATTATATATTCCAGTAT
>tfb1         ------------------------------------------------------------------

>fba-pfp           ------------------------------------------------
>pps               CCG-------------------------------------------------
>orf1155-lrp       ACCACTATAGCGTCC--------AGAATAAAGAAGCTCG-----------
>gar1-tfb2         ACACTTACTCTATGAAGAGGGTCGGCACAACTTTCTATGACGTAATCGGG
>tfb1              ------------------------------------------------
```

FIGURE 10. DNA sequence of the *fba-pfp*, *pps*, *orf1155-lrp*, *gar1-tfb2*, and *tfb1* probes used for analysis of *Ttx*-TFB1 and *Ttx*-TFB2 binding in absence and presence of *Ttx*-TBP. Predicted core promoters (BRE/TATA) are indicated in bold and the start codon of the ORF or the first ORF of the operon is indicated in lower-case bold letters.

Strikingly, TFB2 of *T. tenax* was found to bind stable to the *fba-pfp* DNA probe in the absence of *Ttx*-TBP. In addition, multiple bands are observed, when increasing *Ttx*-TFB2 amounts are added (Figure 11), suggesting that monomers of *Ttx*-TFB2 bind to different sites of the DNA or different monomers bind cooperatively to a specific site. The addition of *Ttx*-TFB2 does not influence *Ttx*-TBP binding to the *fba-pfp* probe. Strikingly, a negative effect on the *Ttx*-TFB2 binding is observed when *Ttx*-TBP is added (Figure 11, lanes 4 and 9). *Ttx*-TFB2 binding in absence or presence of *Ttx*-TBP was also confirmed for other DNA probes (i.e. *tfb1*, *gar1-tfb2*, *pps* and *orf1155-lrp*) by S. KONING in the laboratory of Prof. Dr. BETTINA SIEBERS. To our knowledge, this is the first example known so far, of a TFB homologue, which binds DNA independently of TBP in Archaea.

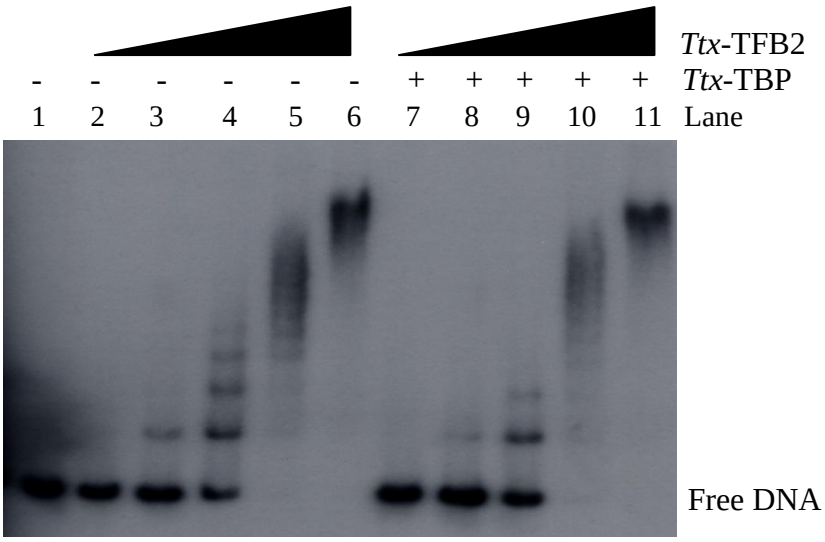


FIGURE 11. DNA binding of *Ttx*-TFB2 in the presence and absence of *Ttx*-TBP.
The DNA probe comprises the transcriptional control region of the *fba-pfp* operon (528 bp, Figure 10). Equal amounts (30 ng) of *Ttx*-TBP and increasing amounts (0, 1.2, 2, 6, 12 and 30 ng) of *Ttx*-TFB2 were added as indicated. Similar results were obtained using the promoter fragment of the genes *tfb1*, *pps*, and operon *orf1155-lrp* as DNA probes.

3.4.2 DNA binding of *Ttx*-TFB1, *Ttx*-TFB2 and *Ttx*-TFB3 using a short-length *fba-pfp* promoter fragment.

Notably, *Ttx*-TFB1 and *Ttx*-TFB2 seem to bind in a different mode to the five different promoter regions tested (see 3.4.1), which regulate transcription of a diverse set of genes encoding for example enzymes or proteins involved in different cellular process, such as, transcription, transcription regulation or metabolism.

In light of this observation, the specificity of the DNA binding of *Ttx*-TFBs was examined in this study using *fba-pfp* promoter fragment as DNA probe. Comparative studies of carbohydrate metabolism in hyperthermophilic Archaea indicate that sugars are generally metabolized using modifications of the classical Entner-Doudoroff (ED) and Embden-Meyerhof-Parnas (EMP) pathway, known from Eukaryotes and Bacteria. Both modifications were demonstrated in *T. tenax* (for recent review see ZAPARTY et al., 2008a). The EMP modification in *T. tenax* is characterized by a non-allosteric reversible PP_i –dependent phosphofructokinase (*pfp* gene) encoded together with the archaeal type class I fructose-1,6-biphosphate aldolase (*fba* gene) in the *fba-pfp* operon. DNA microarray analyses revealed that the transcription of the *fba-pfp* operon is up-regulated under heterotrophic growth conditions (ZAPARTY et al., 2008b). The *fba-pfp* promoter region was been studied in detail, therefore it was selected as DNA probe to define *Ttx*-TFBs binding sites. The transcriptional start position

has been determined previously using primer extension and has been shown to correspond with the translational start position (SIEBERS et al., 2004). The predicted *fba-pfp* promoter is located between positions -38 and -23 (Figure 10) and was identified by comparison with the consensus crenarchaeal promoter predicted by SLUPSKA and co-workers (2001).

In order to examine the specificity of the *Ttx*-TFBs binding, firstly, a shorter fragment including the *fba-pfp* promoter with a length of 60 bp (*fba-pfpWT*-60bp) (from position -50 to +10) was used as DNA probe in EMSAs (Figure 12). The *fba-pfpWT*-60bp DNA probe was prepared by oligonucleotide hybridization (see 2.4.13).

As shown in Figure 12A and B, *Ttx*-TFB1 and *Ttx*-TFB2 bind the *fba-pfp*WT-60bp DNA probe as observed previously for larger fragments (Figure 9 and 11). Therefore, these EMSA demonstrate that *Ttx*-TFB1 is able to form stable complexes with DNA only in presence of *Ttx*-TBP, whereas *Ttx*-TFB2 binds stable to DNA alone and it appears that binding of *Ttx*-TFB2 is not affected by *Ttx*-TBP, because no supershift is detected. As mentioned previously, also a negative effect of *Ttx*-TBP on *Ttx*-TFB2-DNA complex formation is observed when comparing the lanes with equal amount of *Ttx*-TFB2 in absence and presence of *Ttx*-TBP (Figure 12B, lanes 9, 13 and 10, 14). Interestingly, also *Ttx*-TFB3 binds the *fba-pfp*WT-60bp DNA probe in absence of *Ttx*-TBP indicating a TBP independent binding (Figure 12C) similar to *Ttx*-TFB2.

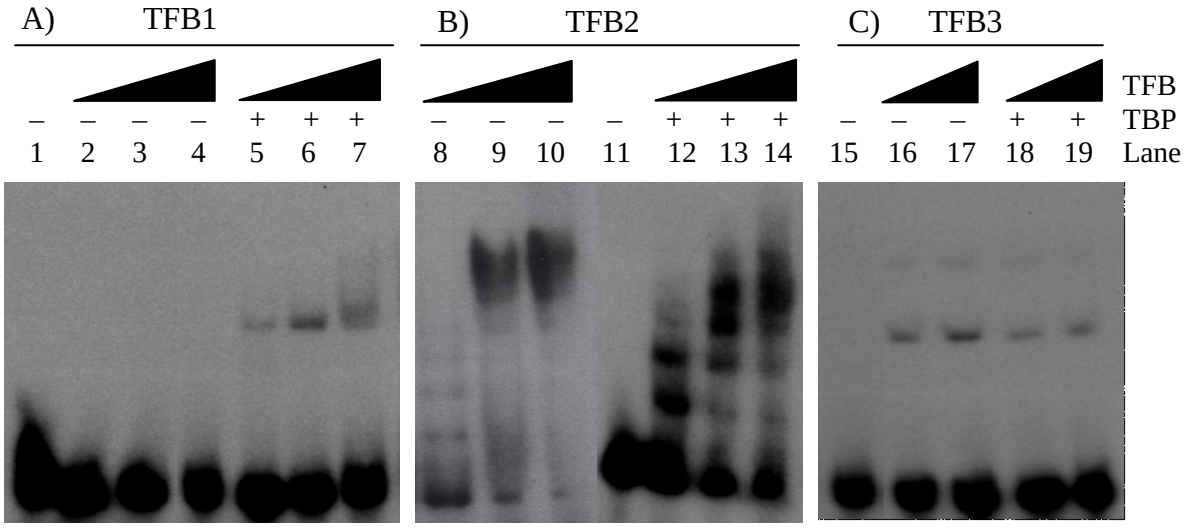


FIGURE 12. DNA Binding of *Ttx*-TFB1, *Ttx*-TFB2 and *Ttx*-TFB3 to the *fba-pfp*WT-60bp probe in the presence and absence of *Ttx*-TBP. The DNA probe comprises the transcriptional control promoter region [(-50) to (+10)] of the *fba-pfp* operon. Equal amounts (30 ng) of *Ttx*-TBP and increasing amounts of *Ttx*-TFB1 (0, 20, 50 and 100 ng), *Ttx*-TFB2 (0, 20, 50 and 100 ng) and *Ttx*-TFB3 (0, 120 and 200 ng) were added as indicated.

3.4.3 Influence of competitor DNA on the DNA binding of *Ttx*-TFB2 to the *fba-pfp*WT-60bp.

In order to study the influence of competitor DNA on the *Ttx*-TFB2 binding to the *fba-pfpWT*-60bp DNA probe, competition experiments were performed. Unlabeled *fba-pfpWT*-60bp DNA probe was used as specific competitor DNA and was added at 2, 5 and 10-fold excess upon the labeled DNA probe. In absence of competitor DNA the formation of the *Ttx*-TFB2 – DNA complex is observed and in the presence of increasing amounts of specific competitor DNA the complex formation is reduced (Figure 13) indicating that the unlabeled competitor DNA (*fba-pfpWT*-60bp) recruits *Ttx*-TFB2. The competition was significally enhanced in the presence of *Ttx*-TBP (Figure 13), supporting the aforementioned negative effect of *ttx*-TBP on *ttx*-TFB2 binding, which has been observed in previous experiments (Figure 11, 12, 13).

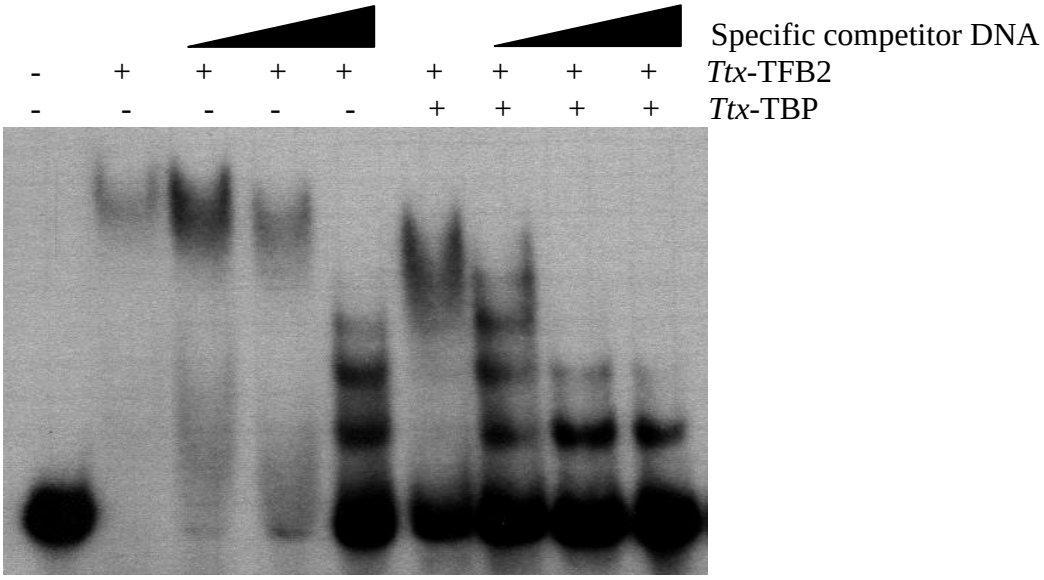


FIGURE 13. DNA binding of *Ttx*-TFB2 to the *fba-pfp*WT-60bp probe in the presence and absence of competitor DNA (unlabeled *fba-pfpWT*-60bp, comp DNA). Equal amounts of *Ttx*-TFB2 (50 ng) and *Ttx*-TBP (30 ng) and increasing amounts (0, 2x, 5x and 10x) of the specific competitor DNA were added as indicated above each lane.

3.4.4 DNA binding of *Ttx*-TFB1 and *Ttx*-TFB2 to mutated *fba-pfp* promoter fragments.

Two mutant sequences of *fba-pfpWT*-265bp sequence were previously constructed via PCR mutagenesis and recombination by S. KONING in the laboratory of Prof. Dr. BETTINA SIEBERS. The derived mutated promoter regions were named *fba-pfpP*-265bp and *fba-pfpB/T*-265bp (Figure 14). In *fba-pfpP*-265bp mutations are located between positions -55 to +15 of the *fba-pfp* promoter region and in *fba-pfpB/T*-265bp between positions -35 to -22 targeting specifically the core promoter (BRE/TATA), respectively (Figure 14).

```
            →M13-R
fba-pfpWT   gagcggataacaatttcacacaggaaacagctatgaccatgattacgccaagcgcgcaattaaccctcact
fba-pfpB/T  gagcggataacaatttcacacaggaaacagctatgaccatgattacgccaagcgcgcaattaaccctcact
fba-pfpP    gagcggataacaatttcacacaggaaacagctatgaccatgattacgccaagcgcgcaattaaccctcact

fba-pfpWT   aaagggaacaaaagctgggtaccgggccccccctcgaggtcgacggtatcgatagtaccgggcccccccctc
fba-pfpB/T  aaagggaacaaaagctgggtaccgggccccccctcgaggtcgacggtatcgatagtaccgggcccccccctc
fba-pfpP    aaagggaacaaaagctgggtaccgggccccccctcgaggtcgacggtatcgatagtaccgggcccccccctc

                                                    -160      -150      -140
fba-pfpWT   gaggtcgacggtatcgataagcttgatatcgaattCCGCCTCAATGGCGGCGTCGGTCGACTCGATATAA
fba-pfpB/T  gaggtcgacggtatcgataagcttgatatcgaattCCGCCTCAATGGCGGCGTCGGTCGACTCGATATAA
fba-pfpP    gaggtcgacggtatcgataagcttgatatcgaattCCGCCTCAATGGCGGCGTCGGTCGACTCGATATAA

            -130      -120      -110      -100      -90       -80       -70
fba-pfpWT   CGATTTAACGCGCATCCAAGAGATATACTTCAATAAAATCATAAAACTCATATCCCATTATATTAATTGC
fba-pfpB/T  CGATTTAACGCGCATCCAAGAGATATACTTCAATAAAATCATAAAACTCATATCCCATTATATTAATTGC
fba-pfpP    CGATTTAACGCGCATCCAAGAGATATACTTCAATAAAATCATAAAACTCATATCCCATTATATTAATTGC

                                                  BRE/TATA          |→
          -60       -50       -40       -30       -20       -10       +1        +10
fba-pfpWT  TCTACTTAATAATATACTTTAGACAAAAAGATATTAAAATGGATAATTGCTCAAGGATCAATGGCAAACCT
fba-pfpB/T TCTACTTAATAATATACTTTAGACA**GCATCGCCGAGAAG**TGGATAATTGCTCAAGGATCAATGGCAAACCT
fba-pfpP   TCTAC**CTCGGCGCAGATGCTATGAAGATAAAGTACACGGGAGATCCCAAGACCTTCTCCTGGGCCGTCAAA**

                   +20       +30       +40       +50       +60          +70      +80
fba-pfpWT   CACCGAGAAATTCTTAAGGATATTCGCGAGGAGGGGGAAGTCCATAATACTGGCCTACGACCACGGCATT
fba-pfpB/T  CACCGAGAAATTCTTAAGGATATTCGCGAGGAGGGGGAAGTCCATAATACTGGCCTACGACCACGGCATT
fba-pfpP    **GTGG**GAGAAATTCTTAAGGATATTCGCGAGGAGGGGGAAGTCCATAATACTGGCCTACGACCACGGCATT

                  +90       +100
fba-pfpWT   GAGCACGGGCCCGCGGACTggatccactagttctagagcggccgccaccgcggtggagctccaattcgccc
fba-pfpB/T  GAGCACGGGCCCGCGGACTggatccactagttctagagcggccgccaccgcggtggagctccaattcgccc
fba-pfpP    GAGCACGGGCCCGCGGACTggatccactagttctagagcggccgccaccgcggtggagctccaattcgccc

fba-pfpWT   tatagtgagtcgtattacgcgcgctcactggccgtcgttttacaacgtcgtgactgggaaaaccctggc
fba-pfpB/T  tatagtgagtcgtattacgcgcgctcactggccgtcgttttacaacgtcgtgactgggaaaaccctggc
fba-pfpP    tatagtgagtcgtattacgcgcgctcactggccgtcgttttacaacgtcgtgactgggaaaaccctggc
                                                                  ←M13-F
```

FIGURE 14. DNA sequence of *fba-pfpWT* (528 bp) and the mutated *fba-pfp* promoter regions (528bp). The *fba-pfpWT*-528bp fragment harbors 265-bp of the *fba-pfp* promoter region (-165 to +100) (in upper-case letters) flanked by 142 bp and 121 bp of *pBlueScript* II KS(+) vector sequence (in lower-case letters). The *fba-pfpB*/T-265bp and *fba-pfpP*-265bp fragments contain mutations at positions from (-35) to (-22) and (-55) to (+15), respectively (bold). The predicted *fba-pfp* promoter is boxed (from (-38) to (-21)) and the determined transcription and translation start site is marked at position (+1, ATG start codon). The universal primers, M13-F and M13-R are underlined.

In the presence of *Ttx*-TBP and *Ttx*-TFB1, binding to both mutated promoter regions (*fba-pfpP*-265bp and *fba-pfpB*/*T*-265bp) is observed as shown in Figure 15A. Unfortunately, the quality of Figure 15A (*fba-pfpB*/*T*-265bp) is low, however, the signals are clearly visible at the respective autoradiogram. As illustrated by the smear observed during electrophoresis, the *Ttx*-TFB1 binding appears to be unstable to both mutated DNA probes in presence and in absence of *Ttx*-TBP. Thus, although the BRE/TATA region of the *fba-pfp* promoter is mutated the *Ttx*-TFB1/*Ttx*-TBP complex is able to bind DNA, suggesting either unspecific binding to the core promoter or unstable binding to alternative up- or downstream sequences.

A)

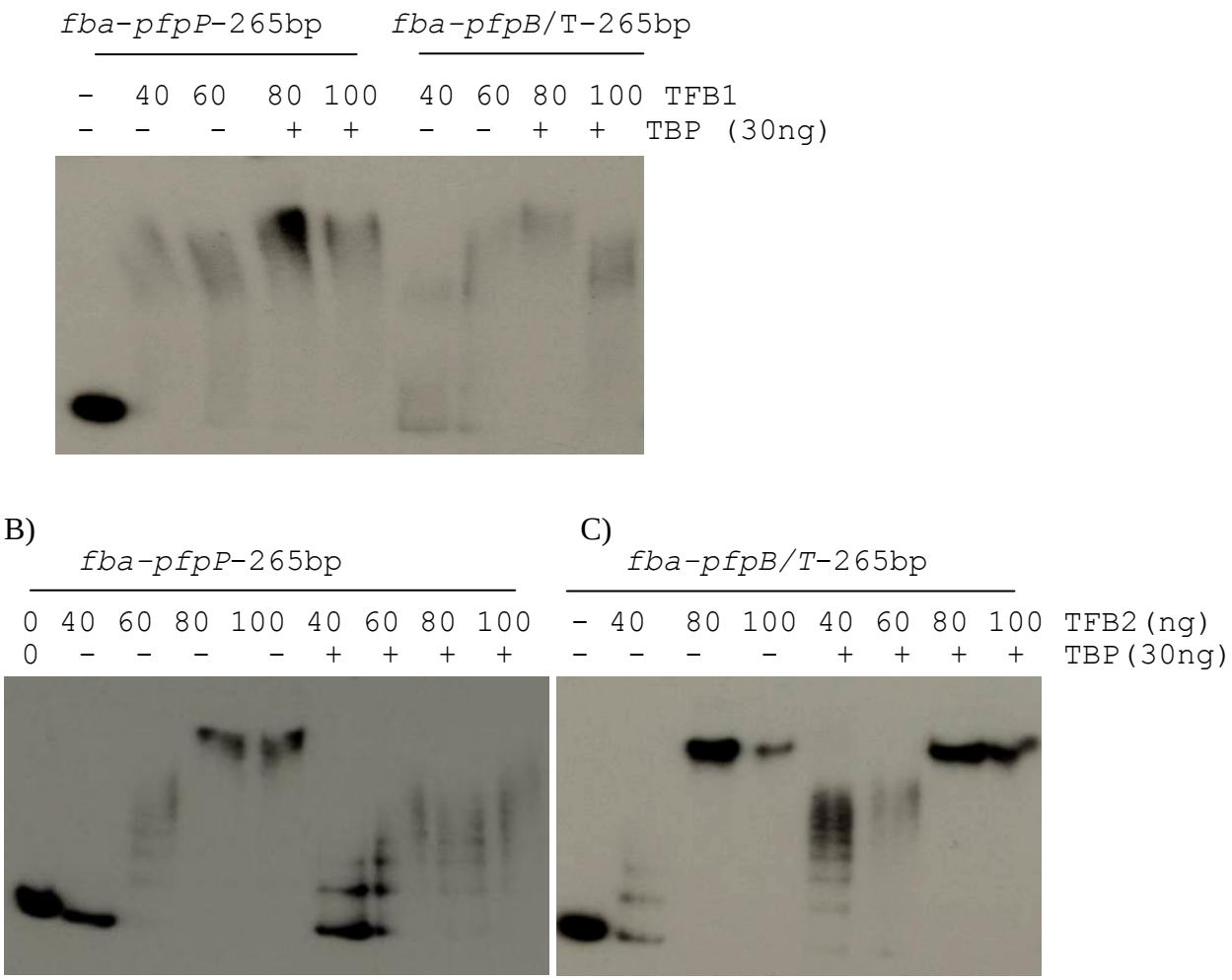


FIGURE 15. DNA binding of *Ttx*-TFB1 (A) and *Ttx*-TFB2 (B) to mutated promoter regions of the *fba-pfp* operon. DNA probes (265 bp) of the mutated promoter regions (*fba-pfpB/T* and *fba-pfpP*) were tested. The protein amounts are indicated above each lane. The sequences of the mutated promoter regions are shown in FIGURE 14.

Ttx-TFB2 binds to the two mutated DNA probes as observed previously for the WT probe (Figure 11). The stability of *Ttx*-TFB2 binding to the *fba-pfp* DNA probe is not affected by the mutations as shown by EMSAs using the two mutated DNA probes, *fba-pfpB/T*-265bp (Figure 15B) and *fba-pfpP*-265bp (Figure 15C). Again, *Ttx*-TFB2 shows *Ttx*-TBP independent DNA binding. These results, using mutated fragments as DNA probes, suggest that *Ttx*-TFB2 binds stable but either unspecifically to BRE/TATA region or to sequences outside of *fba-pfp* core promoter region.

3.4.5 Identification of *Ttx*-TFB1/*Ttx*-TBP and *Ttx*-TFB2 binding sites at the *fba-pfp* promoter region.

To identify potential binding sites of *Ttx*-TFB1 and *Ttx*-TFB2, DNAse I and Exonuclease III footprinting assays were performed in collaboration with PD. Dr. WINFRIED HAUSNER from

The Archaea Center, University of Regensburg (Germany). No qualitative changes were observed in the DNase I footprinting patterns in absence or presence of TFBs, probably because the binding of *Ttx*-TFBs to DNA is not strong enough and the *Ttx*-TFB/DNA complex dissociates upon the addition of DNAse I. Therefore, exonuclease III footprinting assays were performed in this work in order to identify the *Ttx*-TFB1/*Ttx*-TBP and *Ttx*-TFB2 binding sites at the *fba-pfp* promoter region. In this technique, 5'- terminally labeled DNA fragments containing the *fba-pfp* promoter region are incubated in the presence of suitable non-specific competitor DNA with purified recombinant proteins. In this work λ*HindIII* DNA (MEB Fermentas, St. Leon Rot, GER) was used as non-specific competitor DNA in the binding assays. After incubation, protein-complexes are detected by digestion with *E. coli* exonuclease III (Exo III). Specifically bound proteins sterically block the 3' to 5' processive digestion of DNA by Exo III, and create novel termination products, which can be distinguished from natural termination events by comparing the array of digestion products with those created by digestion of the probe (DNA fragments) in the absence of added proteins. Usually, Exo III stops 20 to 30 bp before the site where the DNA-binding protein is bound (Figure 16), therefore, the binding site can not be determined directly.

The 528-bp fragment (*fba-pfpWT*-528bp) was amplified by PCR using plasmid pBS-*fpaWT*-265bp as template DNA and the universal primer set M13-R (labeled) and M13-F (unlabeled) allowing the specific labeling of the coding strand. The *fba-pfpWT*-528bp fragment harbors 265-bp of the *fba-pfp* promoter region (-165 to +100) flanked by 142 bp and 121 bp of *pBlueScript* II KS(+) vector sequence (Figure 14). The PCR fragments were purified via 15% polyacrylamide gel, before they were used as DNA probe for the Exo III footprinting assay.

Figure 17 shows the results of the Exo III footprinting assay in the presence of *Ttx*-TFB1 and *Ttx*-TFB2 using the 5'-terminal labeled coding strand of the *fba-pfpWT*-528bp as DNA probe. The Exo III pattern generated in presence of the *Ttx*-TBP/*Ttx*-TFB1/DNA complex reveals two prominent stop sites at positions -16 and +1 (Figure 17, lane 5) indicating 20 bp downstream protein binding, which corresponds to the *fba-pfp* BRE/TATA region (-38 to -23) (Figure 14). Binding to the *fba-pfp* BRE/TATA site is not detected in the Exo III footprinting pattern generated by *Ttx*-TFB1 (300 ng) in absence of *Ttx*-TBP (Figure 17, lane 2). For *Ttx*-TFB2 (225 - 900 ng) either alone or in the presence of *Ttx*-TBP no binding to the BRE/TATA region is observed (Figure 17, lanes 7 to 9). As expected Exo III footprinting analyses in presence of *Ttx*-TBP (50 ng) alone show no protein binding (Figure 17, lane 3). These results indicate that *Ttx*-TBP stabilizes the binding of *Ttx*-TFB1 to the BRE/TATA region of the *fba-*

pfp promoter (Figure 17, lane 5), however, no binding of *Ttx*-TFB2 is observed independent of the presence of *Ttx*-TBP (Figure 17, lanes 7 to 9).

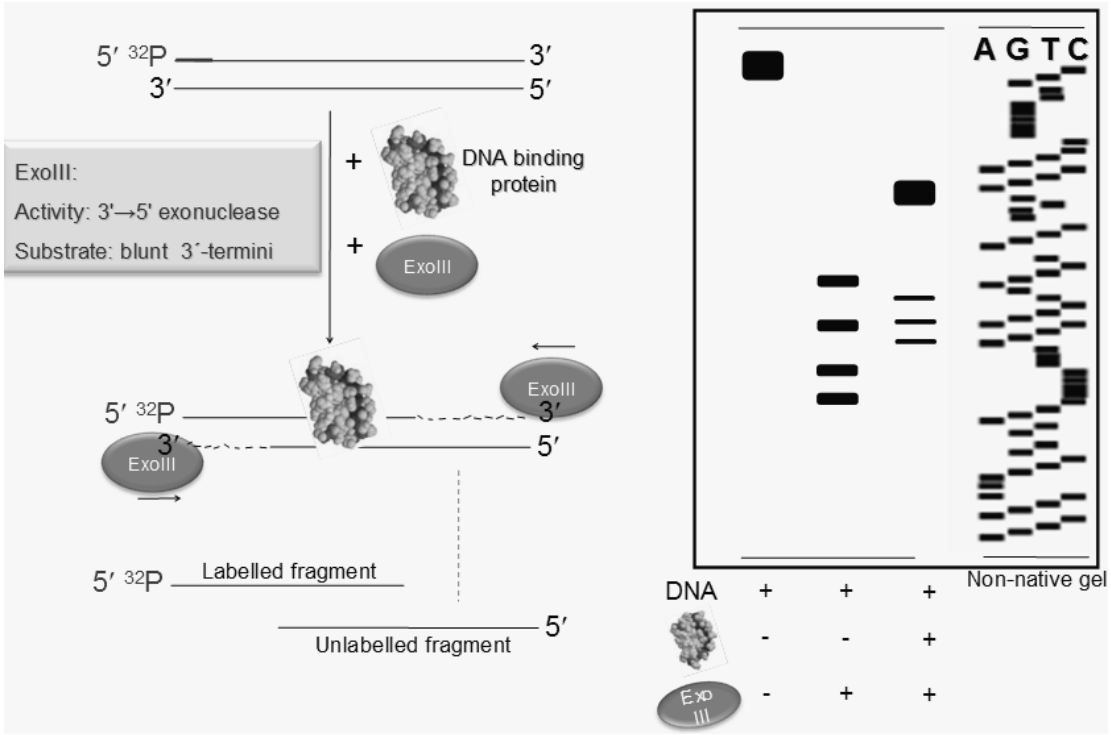


FIGURE 16. Principle of the Exonuclease III footprinting technique.

Strikingly, in the presence of *Ttx*-TFB1 alone several stops of digestion with Exo III were identified in the upstream region of the coding strand of the *fba-pfp* BRE/TATA region (Figure 17, lane 2). The major stops were at positions -76, -58 and -37 (Figure 17, lane 2). A similar Exo III-protection pattern is observed by *Ttx*-TFB2 alone (Figure 17, lanes 7). In absence and in presence of *Ttx*-TBP no difference is detected in the Exo III footprinting pattern for the *Ttx*-TFB2/DNA complex.

The signals corresponding to the above mentioned stops (Figure 17, lane 2 and lane 7) are intensified for *Ttx*-TFB1 (300 ng) when 50 ng *Ttx*-TBP is added (Figure 17, lane 5) and for *Ttx*-TFB2 when increasing protein amounts (225 – 900 ng) are added to the binding assay (Figure 17, lane 7-9). The stops (from position -76 to +1) are highly resistant to longer incubation time with Exo III (Figure 17) and are not detected in the control digestion of DNA without GTFs (Figure 17, lane 1).

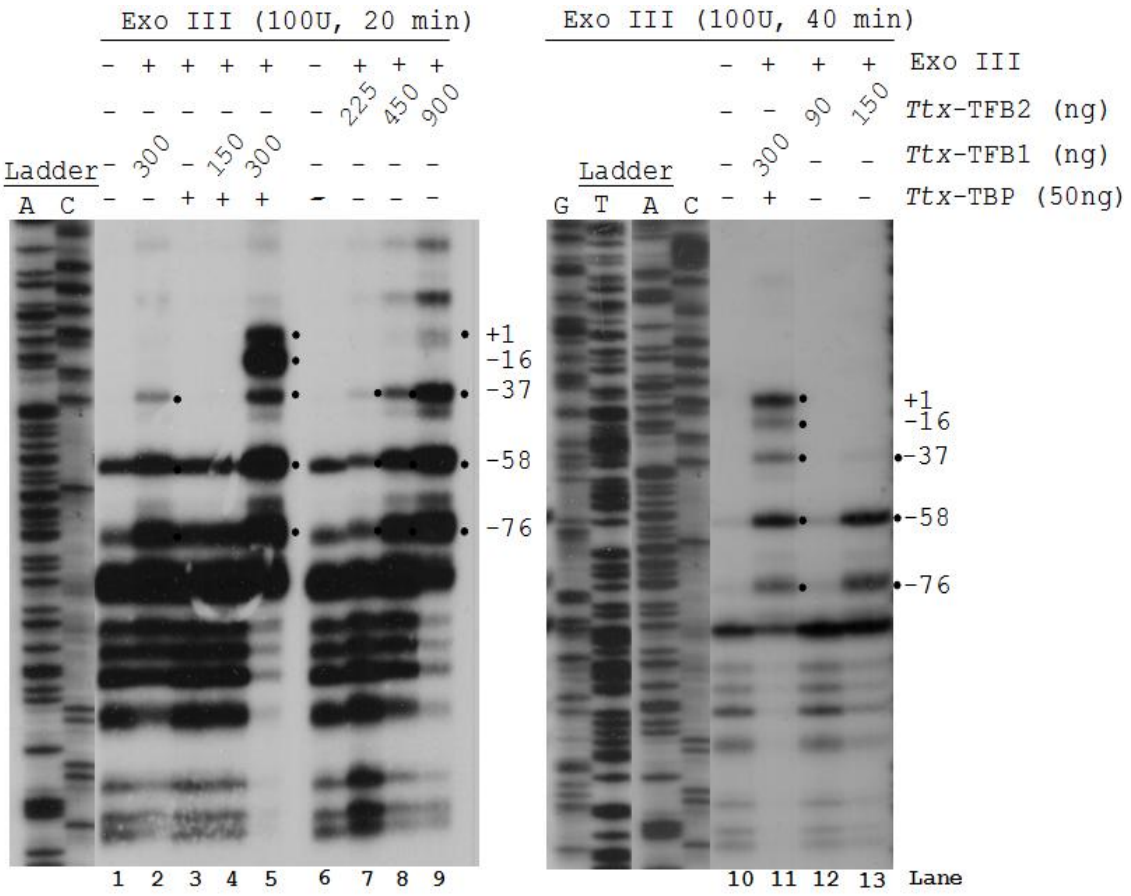


FIGURE 17. Identification of *Ttx*-TFB1 and *Ttx*-TFB2 binding sites at the promoter region [(-165) to +100] of the *fba-pfp* operon from *T. tenax*. Exo III footprinting assay of fragment *fba-pfpWT* (5' end labeled 528 bp DNA fragment of the coding strand). The DNA sequencing reaction (G, T, A and C) of the coding strand labeled at the 5' end is shown. The Exo III stop points caused by bound protein are indicated by dots and are numbered according to the position in respect to the transcription start (+1).

The region located 20 - 30 bp upstream -76/-58 encompasses an AT-rich binding sequence ([-101] 5'-TCAATAAAATCATAAAA-3' [-84]) with similarity to the crenarchaeal consensus promoter (ANAAAA(NNN)CTTTTAAA) (SLUPSKA et al., 2001)) and to the *fba-pfp* promoter ([-38] 5'-ACAAAAAGATATTAAAA-3' [-21]; only 5 bp difference in 17 bp) (Figure 18). Moreover two direct repeats, ATAAA(N2)ATAAA, are identified in this region (-101/-84). It is therefore likely that the -76 to -58 stops are caused by specific binding of *Ttx*-TFB2 to the region -101/-84 located upstream of the *fba-pfp* core promoter.

As mentioned above, EMSAs using mutated *fba-pfp* promoter regions, *fba-pfpP*-265bp and *fba-pfpB*/T-265bp, illustrated that DNA binding of *Ttx*-TFB1 in presence of *Ttx*-TBP is rather unstable but DNA-binding of *Ttx*-TFB2 is stable (Figure 15). The sequences corresponding to the upstream located region (-101/-84) remained as wild-type in both tested mutated promoter regions (*fba-pfpP*-265bp and *fba-pfpB/T*-265bp) (Figure 14). Therefore, the unstable binding of *Ttx*-TFB1 in presence or absence of *Ttx*-TBP to the two mutated promoter region (*fba-*

*pfp*P-265bp and *fba-pfp*B/T-265bp) as well as the stable DNA-binding of *Ttx*-TFB2, which is not affected by mutations in the core promoter region as observed in EMSAs supports a binding of *Ttx*-TFB1 and *Ttx*-TFB2 to the upstream located AT-rich region.

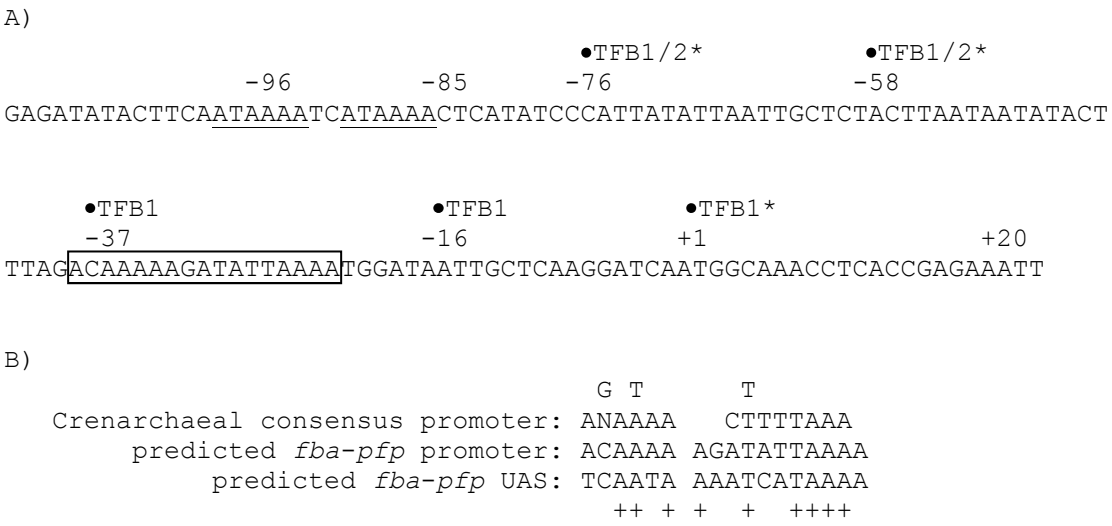


FIGURE 18. Summary of *Ttx*-TFB1/*Ttx*-TBP and *Ttx*-TFB2 -binding sites at the promoter region of the *fba-pfp* operon from *T. tenax* determined by Exo III footprinting assays.
A) Summary of the results of the Exo III footprinting assays shown in Figure 17, which were obtained in several independent experiments. Exo III ususally stops 20 – 30bp before the DNA binding site of the protein. ExoIII stop points caused by protein binding are indicated by dots (●). The most intense stop signals detected (+1, -58 and -76) are marked by asterisk (*).The two direct repeats of the predicted *fba-pfp* upstream activating sequence (UAS) are underlined and the predicted BRE/TATA box of the *fba-pfp* promoter is boxed. B) DNA sequence alignment of the predicted *fba-pfp* promoter, the crenarchaeal consensus promoter and the predicted *fba-pfp* upstream activating sequence (UAS). Matching nucleotides are indicated by (+).

Taken together, results from EMSA and Exo III footprinting analysis confirm that *Ttx*-TFB1 alone binds DNA but the binding appears to be unstable and specific binding to the *fba-pfp* BRE/TATA region is observed only in presence of *Ttx*-TBP. Therefore, the TFB1 paralogue of *T. tenax* behaves like a classical TFIIB factor which shows stable DNA-binding activity only in presence of TBP, which corresponds nicely to the structural and phylogenetic predictions for *Ttx*-TFB1.

The stable binding of *Ttx*-TFB2 to the region -101/-84 suggests that this region might function similar to an upstream activation sequence (UAS) for the *fba-pfp* promoter of *T. tenax*. A detailed inspection of the sequence located between positions -80 and -100 of the different *T. tenax* promoters tested in EMSAs (Figure 10) revealed the presence of AT-rich repeats upstream of the BRE/TATA promoter of other genes (Figure 19). However, differences in promoter architecture are observed, for example, the *pps* gene shows the sequence GTAAT(N5)ATTAC, which comprises and inverted palindromic sequence. Therefore, it will

be important to analyze *Ttx*-TFB1/*Ttx*-TFB2-binding to alternative promoter regions with similar (*pps*) or different (*orf155-lrp*) promoter structure in future experiments.

```
>fba-pfp          (-108) GATATACTTCAATAAAATCATAAAACTCATATCCCATTA (-71)
>pps              (-106) TGAAACTTTAGAGTAATTATATATTACGTGTTTTCTAGG (-68)
>orf1155-lrp   (-108) CGACACCTTGGGCAGTTTGACCGAGGATGCATCTCAGTA (-70)
```

FIGURE 19. DNA sequence of the AT-rich repeats upstream of the BRE/TATA promoter of the *pps* gene and the *fba-pfp* and *orf1155-lrp* operons. AT-rich repeats are indicated in bold.

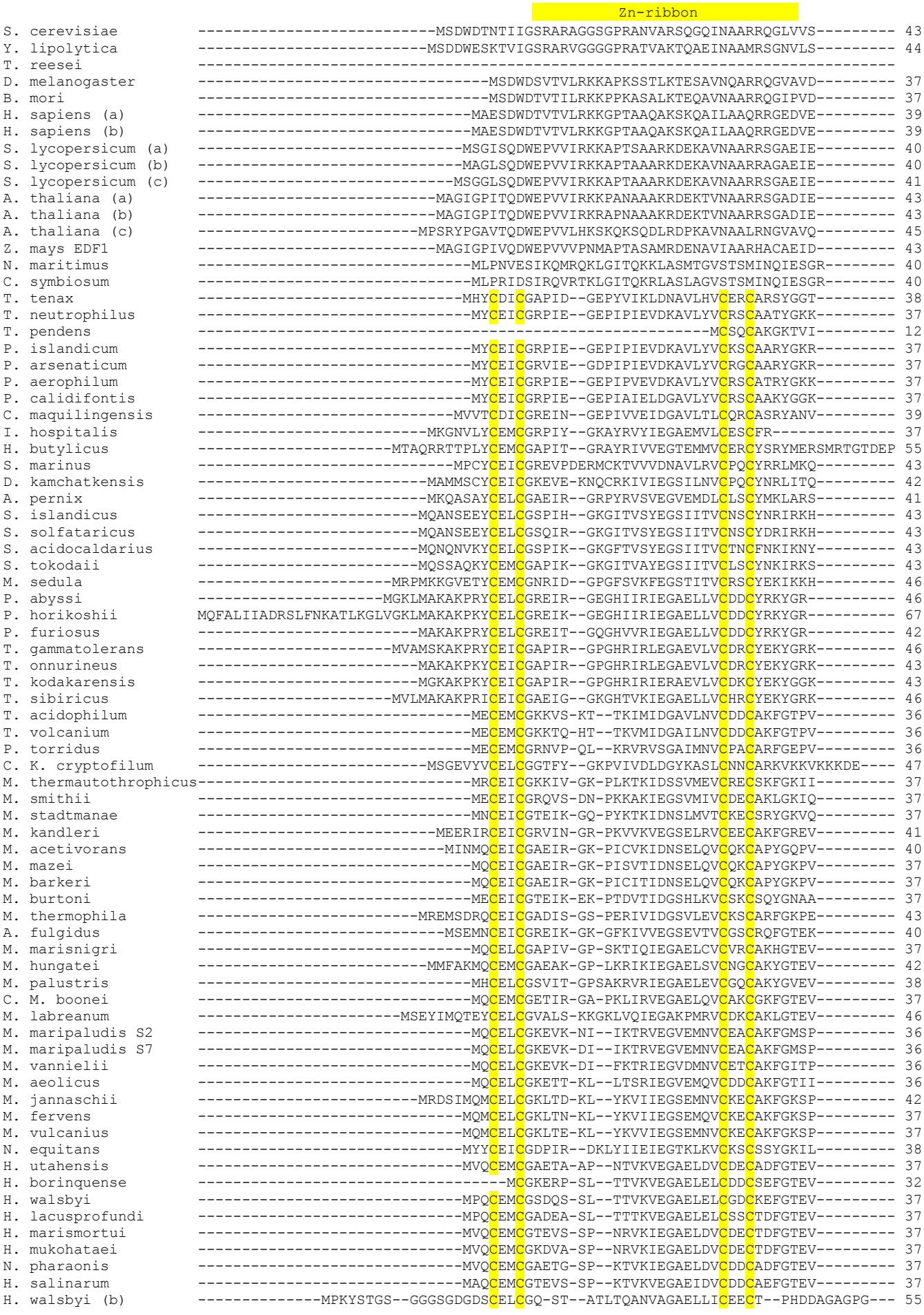
3.5 EUKARYOTIC AND ARCHAEAL MBF1: A COMPARISON.

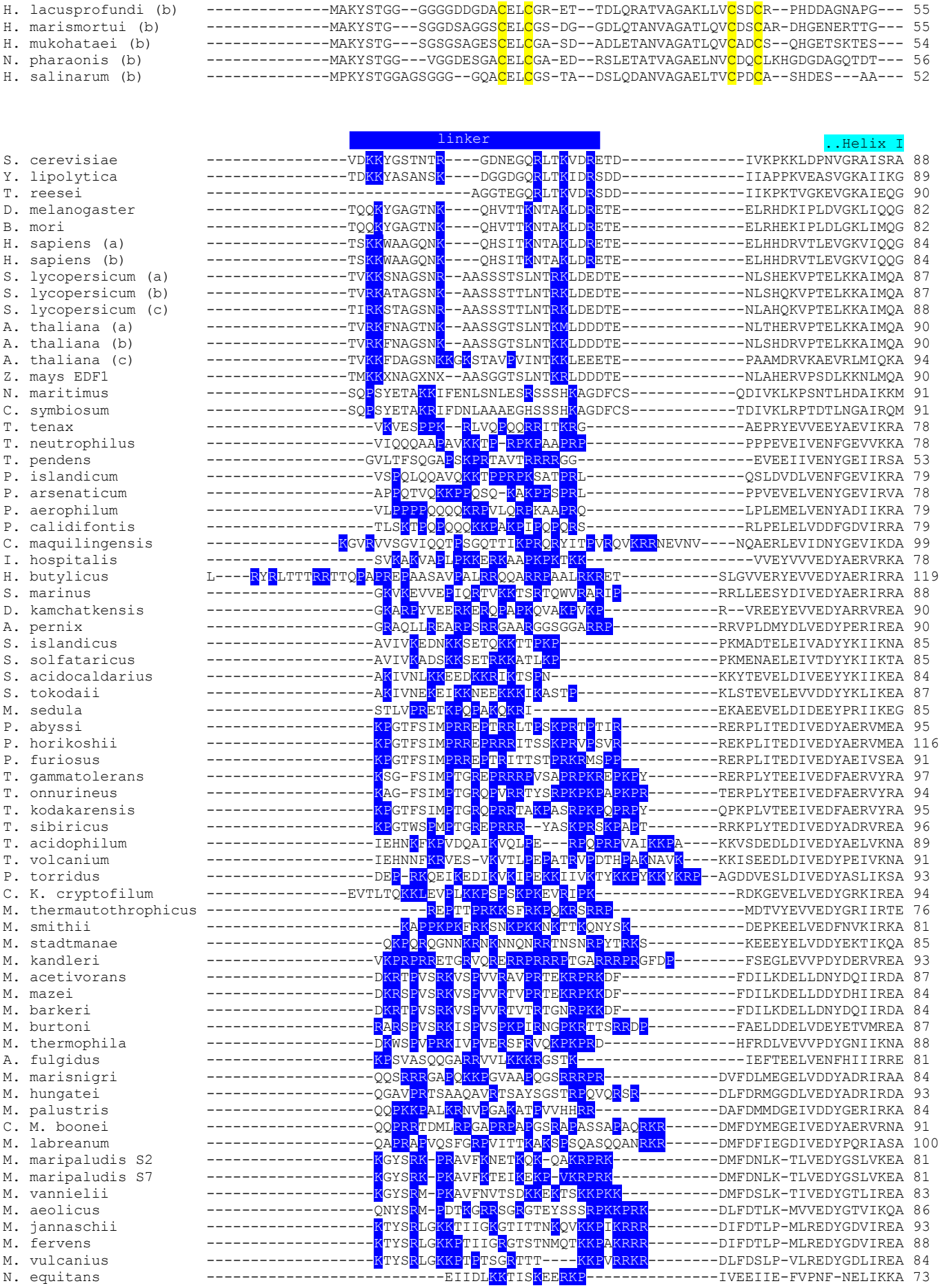
Multiprotein bridging factor 1 (MBF1) is a transcriptional co-activator that mediates transcriptional activation by bridging a sequence-specific activator and TBP in Eukaryotes. Currently no experimental data are available for the biological function of MBF1 in Archaea. Databank searches (BLASTX) revealed sequences with apparent MBF1 similarity in all available 62 archaeal genomes (Reference date: 07.01.2010). All genomes analyzed contain only a single MBF1 homolog with the only exception of Halobacteriales, which harbor two paralogues (KONING et al., 2009). In contrast to previous work, which suggested that *mbf1* is lost in Thaumarcheaota (KONING et al., 2009), a recently proposed new archaeal phylum (BROCHIER-ARMANET et al., 2008), PSI-BLAST searches with *T. tenax* MBF1 revealed sequences with 30% and 27% identity (PSI-BLAST E-value of 10-3) in *Nitrosopumilus maritimus* (genbank accession: ABX12848.1) and *Crenarchaeum symbiosum* (genbank accession: ABK77108.1). In addition, both putative *mbf1* genes are organized in a conserved genome context, which is similar to the one reported previously for Crenarchaeota (e.g. *Thermophilum pendens*) (KONING et al., 2009); therefore, both sequences were included in the phylogenetic analysis.

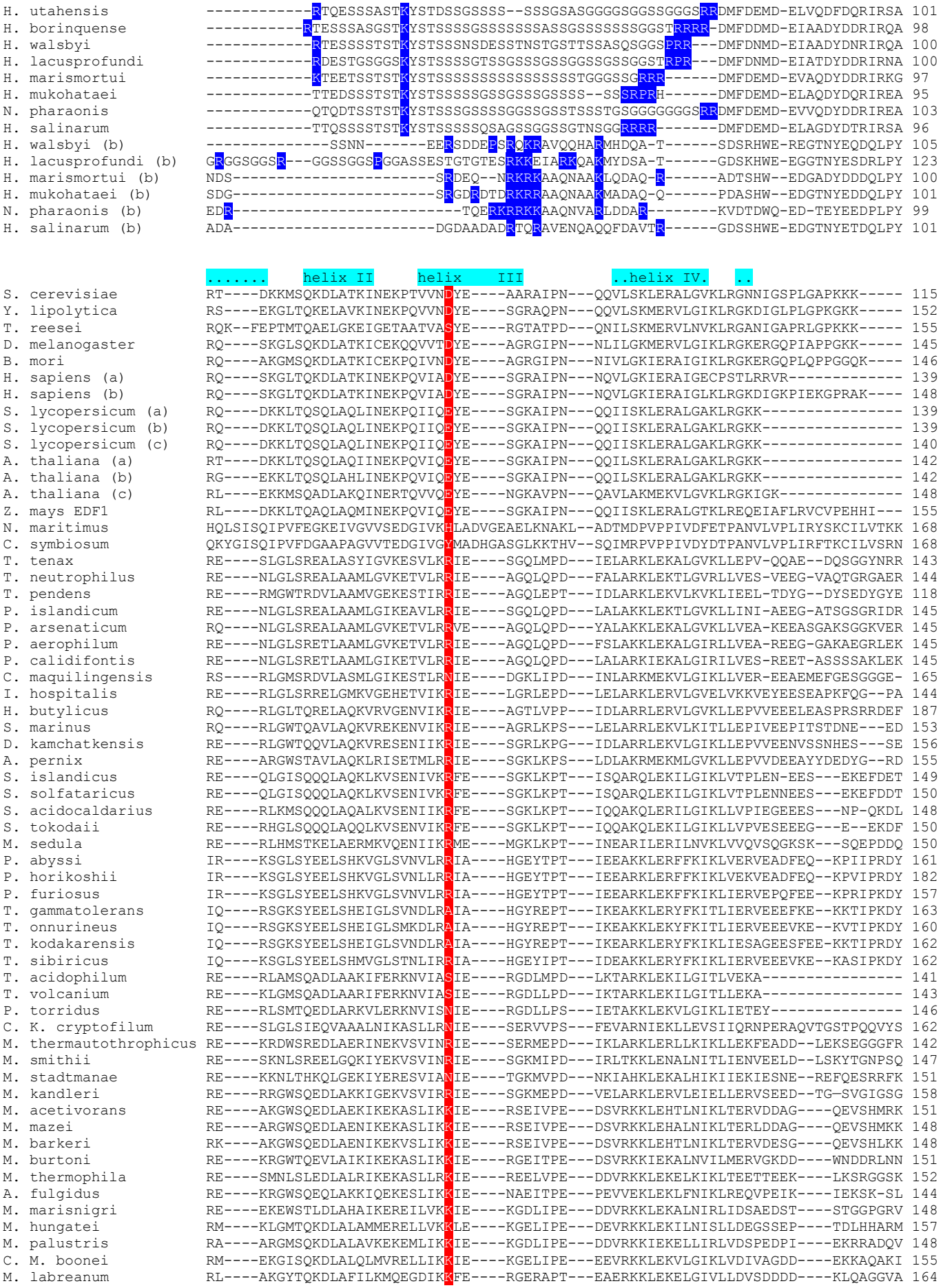
In general, MBF1 comprises an N-terminal domain that is connected by a flexible linker to the C-terminal part, which is composed of a HTH domain and a short C-terminal stretch (Figure 20). In contrast to that, the HTH domain of MBF1 is well-conserved in Archaea and Eukaryotes and comprises four α-helices, the N-terminal domain as well as the C-terminus is more divergent (ARAVIND & KOONIN, 1999; KONING et al., 2009). In the N-terminal part all aMBF1s, with the only exception of *N. maritimus* and *C. symbiosum*, possess a well-conserved Zn-ribbon motif, predicted on the basis of two pairs of cysteine residues, which is absent in eukaryotic counterparts (Figure 20). Due to the presence of the Zn-ribbon a possible direct binding of aMBF1 to DNA has been suggested previously (ARAVIND & KOONIN,

1999; KONING et al., 2009). As shown in Figure 20 all aMBF1, except for Thermoplasmatales and MBF1b paralogues of Halobacteriales, comprise a C-terminal extension with a well-conserved four amino acid motif "[TS]-[LIVMF]-G-[DENI]".

Based on the multiple sequence alignment a phylogenetic tree (distance-based neighbour-joining) was constructed (Figure 21). Bootstrap data support four distinct branches: (I) MBF1 homologues of Eukaryotes, (II) MBF1 homologues of all Archaea, except the cluster grouping the second MBF1 paralogues present in Halobacteriales, which define the third branch (III) and the cluster of Thaumarcheaota (*N. maritimus* and *C. symbiosium*), which define the fourth branch (III). Interestingly, the topology of the unrooted tree suggests two groups for the archaeal branch II, the first one includes MBF1 homologues from Crenarchaeaota, and part of the Euryarchaeota (i.e. Thermococcales and Thermoplasmatales). The second archaeal branch (IIb) comprises MBF1 homologues of Korarchaeota, Nanoarchaeota and all the residual Euryarchaeota. It is noteworthy that most of the archaeal (hyper)thermophiles, are grouped in branch IIa and only few (hyper)thermophiles (*Methanobacter thermoautotrophicus*, *Methanopyrus kandleri*, *Methanosaeta thermophila* and *Archaeaglobus fulgidus*), are found in branch IIb (Figure 21).


Zn-ribbon
S. cerevisiae
Y. lipolytica
T. reesei
D. melanogaster
B. mori
H. sapiens (a)
H. sapiens (b)
S. lycopersicum (a)
S. lycopersicum (b)
S. lycopersicum (c)
A. thaliana (a)
A. thaliana (b)
A. thaliana (c)
Z. mays EDF1
N. maritimus
C. symbiosum
T. tenax
T. neutrophilus
T. pendens
P. islandicum
P. arsenaticum
P. aerophilum
P. calidifontis
C. maquilingensis
I. hospitalis
H. butylicus
S. marinus
D. kamchatkensis
A. pernix
S. islandicus
S. solfataricus
S. acidocaldarius
S. tokodaii
M. sedula
P. abyssi
P. horikoshii
P. furiosus
T. gammatolerans
T. onnurineus
T. kodakarensis
T. sibiricus
T. acidophilum
T. volcanium
P. torridus
C. K. cryptofilum
M. thermautothrophicus
M. smithii
M. stadtmanae
M. kandleri
M. acetivorans
M. mazei
M. barkeri
M. burtoni
M. thermophila
A. fulgidus
M. marisnigri
M. hungatei
M. palustris
C. M. boonei
M. labreanum
M. maripaludis S2
M. maripaludis S7
M. vannielii
M. aeolicus
M. jannaschii
M. fervens
M. vulcanius
N. equitans
H. utahensis
H. borinquense
H. walsbyi
H. lacusprofundi
H. marismortui
H. mukohataei
N. pharaonis
H. salinarum
H. walsbyi (b)


linker
..Helix I


H. utahensis
H. borinquense
H. walsbyi
H. lacusprofundi
H. marismortui
H. mukohataei
N. pharaonis
H. salinarum
H. walsbyi (b)
H. lacusprofundi (b)
H. marismortui (b)
H. mukohataei (b)
N. pharaonis (b)
H. salinarum (b)
helix II
helix III
..helix IV.
S. cerevisiae
Y. lipolytica
T. reesei
D. melanogaster
B. mori
H. sapiens (a)
H. sapiens (b)
S. lycopersicum (a)
S. lycopersicum (b)
S. lycopersicum (c)
A. thaliana (a)
A. thaliana (b)
A. thaliana (c)
Z. mays EDF1
N. maritimus
C. symbiosum
T. tenax
T. neutrophilus
T. pendens
P. islandicum
P. arsenaticum
P. aerophilum
P. calidifontis
C. maquilingensis
I. hospitalis
H. butylicus
S. marinus
D. kamchatkensis
A. pernix
S. islandicus
S. solfataricus
S. acidocaldarius
S. tokodaii
M. sedula
P. abyssi
P. horikoshii
P. furiosus
T. gammatolerans
T. onnurineus
T. kodakarensis
T. sibiricus
T. acidophilum
T. volcanium
P. torridus
C. K. cryptofilum
M. thermautothrophicus
M. smithii
M. stadtmanae
M. kandleri
M. acetivorans
M. mazei
M. barkeri
M. burtoni
M. thermophila
A. fulgidus
M. marisnigri
M. hungatei
M. palustris
C. M. boonei
M. labreanum

```
M. maripaludis S2      RE----KKNMTLEELSRAVGIKESLIHKIE----RNEIEPE---EKYVKILEKALGISFYE--EGDL---NYETSNEDS 144
M. maripaludis S7      RE----KKNMTLEELSRAVGIKESLIHKIE----RNEIEPE---EKYVKILEKALGISFYE--EGDY---NYEANNDES 144
M. vannielii           RE----KRNMTLEGLARTAGIKESLLHKIE----RNEIEPE---EKYVNILERELKISLYE--EGTY---NYESKDENA 146
M. aeolicus            RE----KRGWTLKELAKQIGIKESTLHKIE----RNELEPE---EKYVKRLEKELNITLYE--GSSE---EYEGGADDS 149
M. jannaschii          RE----KRGLSIEELAKKLKMKASTLQKFE----RYELEPN---EKEIKILEKELKISLTESIGEET---SYYGGRDED 158
M. fervens             RE----KRGLSIEDLAKKLKMKASTLQKFE----RYELEPN---EREIKILEKELKINLTESAGEES---PYYAGGDEE 153
M. vulcanius           RE----KMGLSIEELAKKLKMKSSTLQKFE----RYELEPN---EKEIKILEKWLKISLTENVGEDD---QFYSGESDE 149
N. equitans            RE----EKGLTQSDLAKLLHTDINTISKIE----SGDYYPS---EKLAKKIEKLLGIKIMEKKKVALS----QEKAKEE 137
H. utahensis           RE----AADMSQEELADQLNEKASLIRKLE----HGDHLPS---DDVQQKLERALDIELTESGGTDD-DADWDSGSAVG 168
H. borinquense         RE----NNGMSQEDLADSLNEKASLIRKLE----RGDILPP---DNVRKKLERKLDISLVEGGDEEE--SEWSGGSS-T 163
H. walsbyi             RE----SRSLSQEDLADSLNEKASLIRKLE----RSDILPS---DDVREKLERRLDISLVEGTDTDE--DEWSGGSS-T 165
H. lacusprofundi       RE----SRGLSQEELADQLNEKASLIRKLE----RGDTLPT---DDIQRKLESELDITLVEGESADD--ADWDSGDA-G 165
H. marismortui         RE----SQGLSQEELAKQLNEKASLIRKLE----QGNSLPS---DDVQKKLESALEISLS-AGGSAD-ETEWSGGSSDG 163
H. mukohataei          RE----ERGLTQEELAGELNLKASLIRKLE----HGDTLPS---DDVQTTLERELDISLS-AGSTDA-DEEWSGGSSSG 161
N. pharaonis           RE----AASMTQEELANSLNEKASLIRKLE----RGEVLPS---DSVQRKLERELDISLS-TGGGSD-DNDWSGGSSTG 169
H. salinarum           RE----QAGLSQEELADELNEKASVIRKLE----HGDSLPS---DDVREELEDYLGILLTESGAEDASDTDWSSGGDSA 164
H. walsbyi (b)         LV----SEYGEIAASARQAAGFT----IDE----LATELEVESGDILAVEQGRATRAGVGGSLIRELESTLDITLVDE- 161
H. lacusprofundi (b)   LV----SGYGDDVAAARQDAGLT----VEE----LAEELDVDEDDLFAVEDGRAATAGVGGSVVRALEERLGVDIVDE- 179
H. marismortui (b)     LV----SKYGERVTEARQDEGLQ----TSE----LAEELDLDDADILAVEQGRATQANVGGSTIKALEQYLDIDLVESN 156
H. mukohataei (b)      LV----SDYGSVVTEARQDAGLQ----TQE----LAEELSVDESDVLAVEQGRATQAGVGGSVVAKLEDFLDVELAE-- 157
N. pharaonis (b)       LV----SDYGTLVEQARQAEGLQ----IDE----LAREVGADEDDVVAVEQGRAARANVGGSLISALEERLDIELADT- 155
H. salinarum (b)       LV----SDYGERVVRARQDAGLQ----RSE----LAAELEIDDADVLAVEQARATKANVGGSVIAALEDFLDVQLSDD- 157

S. cerevisiae          ------------------ 151
Y. lipolytica          ------------------ 152
T. reesei              ------------------ 155
D. melanogaster        ------------------ 145
B. mori                ------------------ 146
H. sapiens (a)         ------------------ 139
H. sapiens (b)         ------------------ 148
S. lycopersicum (a)    ------------------ 139
S. lycopersicum (b)    ------------------ 139
S. lycopersicum (c)    ------------------ 140
A. thaliana (a)        ------------------ 142
A. thaliana (b)        ------------------ 142
A. thaliana (c)        ------------------ 148
Z. mays EDF1           ------------------ 155
N. maritimus           SKIMGIITASDTLKMME- 185
C. symbiosum           SEIVGIVTALDTLKMME- 185
T. tenax               ELTLGDVAELRDEE---- 157
T. neutrophilus        GLTLGEVAEIRDGGEE-- 160
T. pendens             GVTLGDIAEFRDRGDI-- 134
P. islandicum          GLTLGEVAEIRDSDEK-- 161
P. arsenaticum         GLTLGEVVEIREDGEK-- 161
P. aerophilum          GLTLGEIAEIREDGEK-- 161
P. calidifontis        GLTLGEVAEIRDEGEE-- 161
C. maquilingensis      -VTLGDVVEIRRKDKEGQ 182
I. hospitalis          ELTLGDVAVLRKE----- 157
H. butylicus           YLTLGDIAEIRED----- 200
S. marinus             YYTIGDFIKIKKKK---- 166
D. kamchatkensis       DLTIGDLIRFKRE----- 171
A. pernix              YITLGDIVVVDRDEE--- 169
S. islandicus          ELTLGDVVNIKEGKK--- 164
S. solfataricus        GLTLGDVVNIKEGKK--- 165
S. acidocaldarius      GLTLGDIVNIREGKK--- 163
S. tokodaii            ELTLGDVANIREGKK--- 165
M. sedula              TLTLGDIIRIREGKK--- 165
P. abyssi              EPTLGDIARIKIRKKKKK 179
P. horikoshii          EPTLGDIARIKVKKKKKK 200
P. furiosus            EPTLGDIARIKVKKRKK- 174
T. gammatolerans       EPTLGDIANIRIRKRKK- 180
T. onnurineus          EPTLGDIANIKIRKRKK- 177
T. kodakarensis        EPTLGDIANIRIKKRKKK 180
T. sibiricus           EPTLGDIANIRVKKKKK- 179
T. acidophilum         ---------------- 141
T. volcanium           ---------------- 143
P. torridus            ---------------- 146
C. K. cryptofilum      SVTLGEAVEVKRKRRKR- 179
M. thermautothrophicus GATIGDIARIKRG----- 155
M. smithii             GRTLGNVVKIKKR----- 160
M. stadtmanae          EATLGDIARIKRK----- 164
M. kandleri            ELTLGDVVEIRKK----- 171
M. acetivorans         DTTLGDIVKIKRK----- 164
M. mazei               DTTLGDIVKIKRK----- 161
M. barkeri             DMTLGDIVKIKRK----- 161
```

```
M. burtoni              GTTLGDIVTIKKK----- 164
M. thermophila          VLTLGDIANIRKR----- 165
A. fulgidus             VPTLGDVVVVKRKKK--- 159
M. marisnigri           TMTVGDVISFKKSRK--- 163
M. hungatei             TTTMGDVIQIKKAKK--- 172
M. palustris            KTTFGDLISIKRQP---- 162
C. M. boonei            TQTLGDLTIIRKAKK--- 170
M. labreanum            STTLGDVLQVKRK----- 177
M. maripaludis S2       EFTLGDFIKVKKRK---- 158
M. maripaludis S7       DFTLGDFIKVKNRK---- 158
M. vannielii            DFTLGDFVKIKKR----- 159
M. aeolicus             EFTLGDMIKIKRK----- 162
M. jannaschii           GFTLGDFIKIKK------ 170
M. fervens              GFTLGDFIKIKK------ 165
M. vulcanius            GFTLGDFIKIKR------ 161
N. equitans             LYSLGDIVELD------- 148
H. utahensis            EYTLGDVVERKDS----- 181
H. borinquense          TTTLGDVVKRKD------ 175
H. walsbyi              TTTLGDVVKRKD------ 177
H. lacusprofundi        TMTLGDVVKRKD------ 177
H. marismortui          EYTLGDVVKRKD------ 175
H. mukohataei           EYTLGDVVKRKD------ 173
N. pharaonis            KTTLGDVVKRKD------ 181
H. salinarum            GLTLGDKVRRKSDDS--- 179
H. walsbyi (b)          ------------------ 171
H. lacusprofundi (b)    ------------------ 189
H. marismortui (b)      ------------------ 167
H. mukohataei (b)       ------------------ 166
N. pharaonis (b)        ------------------ 165
H. salinarum (b)        ------------------ 167
```

FIGURE 20. Multiple sequence alignment of archaeal and eukaryotic MBF1s.
The alignment was performed using the ClustalX. Dashes indicate gaps in the amino acid sequence introduced to optimize the alignment. Zn-ribbon and linker are depicted by yellow and blue boxes, respectively. The Cys residues that bind the zinc atom in the Zn-ribbon motif are highlighted in yellow and the basic residues (R or K) of the linker are highlighted in blue. Pred-TMBF1: predicted secondary structure of the HTH domain [α-helices] of MBF1 from *T. tenax* performed by the program HHpred (SÖDING, 1995) based on the determined secondary structure of the HTH domain (1x57_A, Probab = 99,48, E-value = 3,5e-14, Indentities = 30%) of human EDF1. The four helices (I-IV) of the HTH domain are indicated by a light blue background. The aspartic acid at position 112 (D112) of yMBF1 (which interacts with asparagine at position 68 (Q68) of yTBP) and its analogous residues in Eukaryotes and Archaea are highlighted in red. The conserved motif, T(S,I)-L(V,M,F,I)-G-D(E,N,I), in the C-terminal extension of aMBF1 is indicated by a green background.

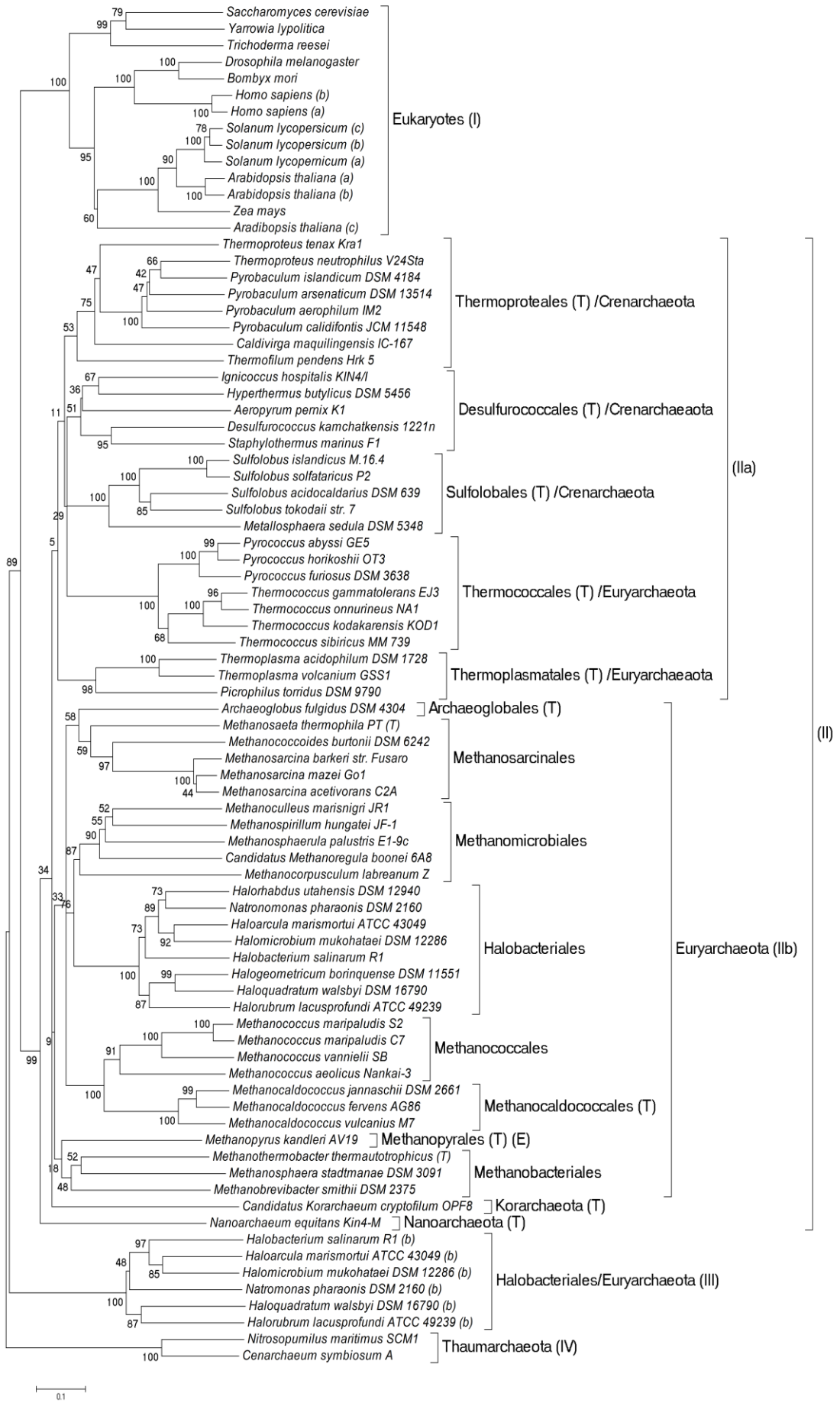

Saccharomyces cerevisiae
Yarrowia lypolitica
Trichoderma reesei
Drosophila melanogaster
Bombyx mori
Homo sapiens (b)
Homo sapiens (a)
Solanum lycopersicum (c)
Solanum lycopersicum (b)
Solanum lycopernicum (a)
Arabidopsis thaliana (a)
Arabidopsis thaliana (b)
Zea mays
Aradibopsis thaliana (c)
Eukaryotes (I)
Thermoproteus tenax Kra1
Thermoproteus neutrophilus V24Sta
Pyrobaculum islandicum DSM 4184
Pyrobaculum arsenaticum DSM 13514
Pyrobaculum aerophilum IM2
Pyrobaculum calidifontis JCM 11548
Caldivirga maquilingensis IC-167
Thermofilum pendens Hrk 5
Thermoproteales (T) /Crenarchaeota
Ignicoccus hospitalis KIN4/I
Hyperthermus butylicus DSM 5456
Aeropyrum pernix K1
Desulfurococcus kamchatkensis 1221n
Staphylothermus marinus F1
Desulfurococcales (T) /Crenarchaeaota
Sulfolobus islandicus M.16.4
Sulfolobus solfataricus P2
Sulfolobus acidocaldarius DSM 639
Sulfolobus tokodaii str. 7
Metallosphaera sedula DSM 5348
Sulfolobales (T) /Crenarchaeota
Pyrococcus abyssi GE5
Pyrococcus horikoshii OT3
Pyrococcus furiosus DSM 3638
Thermococcus gammatolerans EJ3
Thermococcus onnurineus NA1
Thermococcus kodakarensis KOD1
Thermococcus sibiricus MM 739
Thermococcales (T) /Euryarchaeota
Thermoplasma acidophilum DSM 1728
Thermoplasma volcanium GSS1
Picrophilus torridus DSM 9790
Thermoplasmatales (T) /Euryarchaeaota
(IIa)
Archaeoglobus fulgidus DSM 4304
Archaeoglobales (T)
Methanosaeta thermophila PT (T)
Methanococcoides burtonii DSM 6242
Methanosarcina barkeri str. Fusaro
Methanosarcina mazei Go1
Methanosarcina acetivorans C2A
Methanosarcinales
Methanoculleus marisnigri JR1
Methanospirillum hungatei JF-1
Methanosphaerula palustris E1-9c
Candidatus Methanoregula boonei 6A8
Methanocorpusculum labreanum Z
Methanomicrobiales
Halorhabdus utahensis DSM 12940
Natronomonas pharaonis DSM 2160
Haloarcula marismortui ATCC 43049
Halomicrobium mukohataei DSM 12286
Halobacterium salinarum R1
Halogeometricum borinquense DSM 11551
Haloquadratum walsbyi DSM 16790
Halorubrum lacusprofundi ATCC 49239
Halobacteriales
Methanococcus maripaludis S2
Methanococcus maripaludis C7
Methanococcus vannielii SB
Methanococcus aeolicus Nankai-3
Methanococcales
Methanocaldococcus jannaschii DSM 2661
Methanocaldococcus fervens AG86
Methanocaldococcus vulcanius M7
Methanocaldococcales (T)
Methanopyrus kandleri AV19
Methanopyrales (T) (E)
Methanothermobacter thermautotrophicus (T)
Methanosphaera stadtmanae DSM 3091
Methanobrevibacter smithii DSM 2375
Methanobacteriales
Euryarchaeota (IIb)
Candidatus Korarchaeum cryptofilum OPF8
Korarchaeota (T)
Nanoarchaeum equitans Kin4-M
Nanoarchaeota (T)
(II)
Halobacterium salinarum R1 (b)
Haloarcula marismortui ATCC 43049 (b)
Halomicrobium mukohataei DSM 12286 (b)
Natromonas pharaonis DSM 2160 (b)
Haloquadratum walsbyi DSM 16790 (b)
Halorubrum lacusprofundi ATCC 49239 (b)
Halobacteriales/Euryarchaeota (III)
Nitrosopumilus maritimus SCM1
Cenarchaeum symbiosum A
Thaumarchaeota (IV)
0.1

FIGURE 21. Phylogenetic tree of archaeal and eukaryotic MBF1s.
The tree was constructed based on pairwise distance estimates of the expected number of amino acid replacements per site using MEGA4 software (TAMURA et al., 2007). Bar = 0.20 amino acid replacement per site. The bootstrap values are indicated to the left of the branches. Branch length indicates the relative evolutionary distances. (T): (hyper)thermophile, (a, b or c): different homologues in a particular organism.

3.5.1 Heterologous expression and purification of the *T. tenax* MBF1.

The *mbf1* gene (474 bp) of *T. tenax* was first cloned into the subcloning vector *pBlueScript* II KS(+) via PCR mutagenesis and standard cloning techniques. The sequence was confirmed by sequencing of both strands (AGOWA). Afterwards the gene was excised and cloned into the expression vector, pET302 (Table 1). For recombinant expression pET302-*mbf1* was used to transform *E. coli* Rosetta(DE3). The recombinant protein was successfully expressed in the soluble fraction and was enriched from crude extract by heat precipitation (15 min at 80°C). The determined molecular mass (22 kDa) after SDS-PAGE approximately corresponds to the calculated molecular weight of 6xHis-*Ttx*-MBF1 (22.59 KDa, 171 aa). For further biochemical studies, the recombinant protein was purified to apparent homogeneity from crude extract using affinity chromatography and gel filtration (Figure 22).

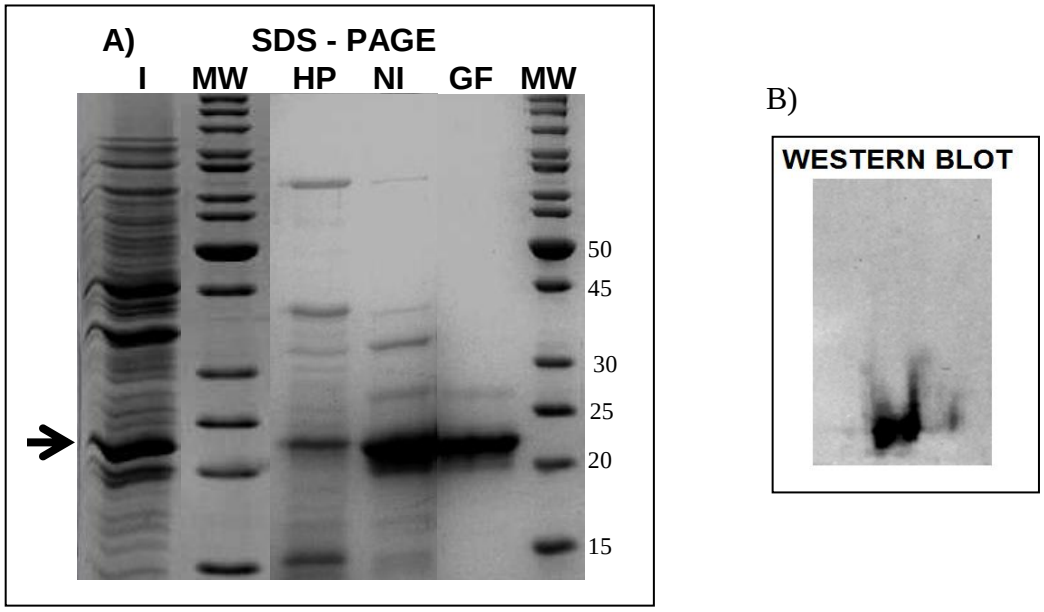


FIGURE 22. SDS-PAGE analysis of the purification of the recombinant 6xHis-*Ttx*-MBF1. Samples after induction with IPTG (I), heat precipitation (15 min, 80°C) (HP), Ni-NTA column (Ni) and after gel filtration (GF) were separated alongside to the molecular weight marker (MW) via SDS-PAGE (12 %). The arrow indicates the position of 6xHis-*Ttx*MBF1. The determined molecular weight is 22 kDa. (B) Detection of the recombinant protein by western blot. 6xHis-*Ttx*-MBF1 was detected using the *Ttx*-MBF1 polyclonal antibody (1:500) generated in this study. Mouse Rabbit-IgG monoclonal antibody conjugated to alkaline phosphatase was used as secondary antibody (1:10,000).

1 mg of the purified 6xHis-*Ttx*-MBF1 was used as antigen for the production of polyclonal antibody (Eurogentec). *Ttx*-MBF1 was specifically detected by western blotting using the respective polyclonal antibody (Figure 22B).

3.5.2 Complementation of yeast MBF1 by archaeal MBF1.

Transcription regulation in Archaea is not extensively studied. Although the archaeal transcription apparatus is similar to the eukaryotic one; most archaeal regulatory transcription factors identified so far are of the bacterial-type (ARAVIND & KOONIN, 1999, LIU et al., 2007). The function of co-activators from Archaea has not been studied so far and the biological role of archaeal MBF1 (aMBF1) as a multiprotein bridging factor has never been documented.

A key question concerning aMBF1, is its *in vivo* function. To address this issue, a complementation study of a *MBF1*-deficient yeast strain with aMBF1 was performed in collaboration with Prof. Dr. ANN EHRENHOFER-MURRAY (Department for Genetics, Faculty of Biology and Geography, University Duisburg-Essen). In this study it was tested whether MBF1s from the hyperthermophile *T. tenax*, and the mesophile *Methanosarcina mazei*, are functional for complementation of a yeast mutant lacking MBF1.

To construct pyMBF1, the 1667-bp *Eco*RI-*Not*I genomic fragment encompassing the entire *MBF1* regulatory and coding regions was cloned into the yeast-bacteria shuttle vector *pRS*316. Afterwards the yMBF1 coding region on the plasmid pyMBF1 was replaced by the *Ttx*-MBF1 and *Mma*-MBF1 coding region using the recombination/gap repair cloning technique in yeast (see 2.8.5). The obtained plasmids p*TMBF1* and p*MMBF1* encode for MBF1 from *T. tenax* and *M. mazei*, under the control of the genomic regulatory region of yMBF1. The wild-type yeast strains, AEY3087, and the deletion strain 45E11 *mbf1*Δ strain (*mbf1*Δ) were kindly provided by Prof. Dr. ANN EHRENHOFER-MURRAY.

In yeast MBF1 is involved in the co-activation of the transcription of *HIS3* gene by bridging TBP and GCN4. The *HIS3* gene encodes imidazole glycerol-phosphate dehydratase, the third enzyme of the histidine biosynthesis pathway. Aminotriazole (AT) is a competitive inhibitor of the *HIS3* gene product. Deletions either of MBF1 or GCN4 in yeast are viable, but sensitive to AT in histidine-free medium. Therefore, complementation by eukaryal MBF1 could be easily detected by growth in the

presence of AT, as demonstrated previously for the three *A. thaliana* MBF1s (TSUDA et al., 2004).

The expression plasmids encoding T*MBF1* and M*MBF1* were introduced into the *mbf1*Δ strain; and the AT sensitivity was compared to that of the wild-type strain transformed with empty plasmid (negative control) and the *mbf1*Δ strain expressing p*yMBF1* (positive control) (Figure 23A). As expected, the negative control, *mbf1*Δ strain, was sensitive to 3-aminotriazole (AT) and AT resistance was restored in the positive control (*mbf1*Δ strain expressing plasmid p*yMBF1*) (Figure 25A). The growth of *mbf1*Δ strain in presence of AT could not be restored by introducing the expression plasmid comprising full-length a*MBF1* genes from *M. mazei* or *T. tenax*, respectively (Figure 23). Western blot analysis of yeast cell extracts, using polyclonal antibody generated against TMBF1, confirmed that TMBF1 was successfully expressed in *mbf1*Δ strain (Figure 23A).

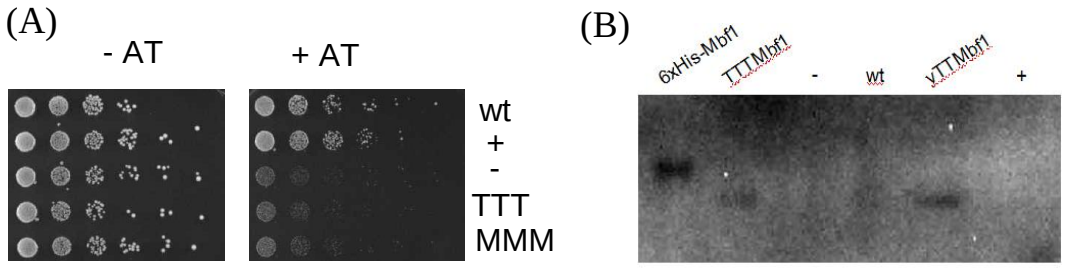


FIGURE 23. Complementation studies of yeast by archaeal MBF1 from *T. tenax* (T) and *M. mazei* (M). (A) AT sensitivity of yeast wild type (WT), yeast *mbf1*Δ strain transformed with py*MBF1* (+), empty vector (-) and aMBF1 from *T. tenax* (pT*MBF1*; TTT) and from *M .mazei* (pM*MBF1*; MMM). After 3 days at 30°C, the cells were diluted to an optical density at 600 nm (OD_{600}) of 0.3, and 10-fold serial dilutions thereof were transferred to YM plates (+Leu, +Met) either in the presence (+AT) or absence (-AT) of 3 mM aminotriazole. The plates were incubated at 30°C for 3 days. (B) The expression of TMBF1 protein in the y*mbf1*Δ strain was detected by western blot using polyclonal antibody against TMBF1.

3.5.3 Chimeric yeast – Archaea MBF1 variants.

Studying protein function by interchanging protein domains is a potentially powerful approach that exploits the natural variability of protein structure. MBF1 from yeast exhibits high structural similarity with *T. tenax* and *M. mazei* MBF1, despite the difference in amino acid sequence; TMBF1 and MMBF1 share 39% and 36% amino acid identity with their yeast counterpart, respectively.

The observed missing complementation of aMBF1 in yeast might be caused by the presence of the Zn-ribbon motif in the N-terminus of aMBF1 or by changes of important amino acids necessary for TBP and GCN4 interaction in aMBF1. To test this possibility and to study the evolutionary conservation of MBF1 domains in Eukaryotes and Archaea, in total 12 chimeric interspecies MBF1 variants were constructed. Each chimeric protein is composed of different combinations of the N-terminal domain, the core domain (including a flexible linker and HTH-domain) and the C-terminal extension from either the hyperthermophile *T. tenax* or the mesophile *M. mazei*, and yeast (Figure 24).

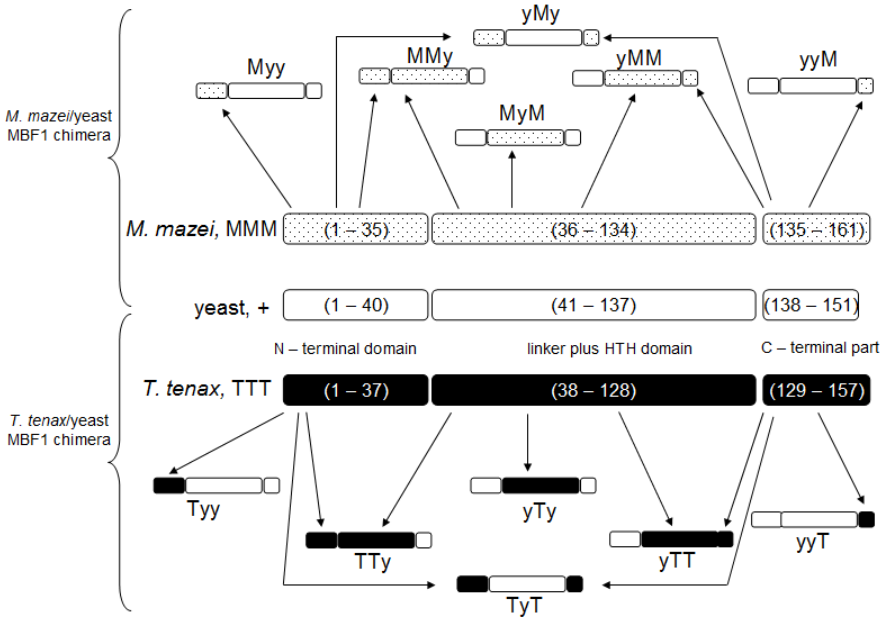


FIGURE 24. Schematic overview of the MBF1 protein from yeast (y), *T. tenax* (T) and *M. mazei* (M), and the Archaea/yeast MBF1chimera generated by the use of recombination/gap repair cloning technique.

The hybrid proteins were constructed using the recombination/gap repair cloning technique in yeast and AT sensitivity was examined (Figure 25). No complementation was observed for any of the aforementioned chimeric MBF1 variants targeting the N-terminal domain or the core domain (i.e. Tyy and Myy; yTy and yMy; TTy and MMy, yTT and yMM; TyT and MyM), indicating that neither the archaeal N-terminal nor the conserved HTH- domain is sufficient to allow for functional MBF1 in yeast.

Next it was studied, whether the well conserved, archaeal C-terminal extension has an influence on yMBF1 function. Interestingly, *mbf1*Δ expressing the chimeric proteins, yyT- and yyM-MBF1, showed similar AT sensitivity as the wild-type and the positive

control (py*MBF1*, Figure 25). However, also the Δ138-151 yMBF1 mutant (yyΔCt control) lacking the yeast C-terminal part (residues 138 - 151) restored WT activity, i.e. AT sensitivity (Figure 25). This observation clearly indicates that the archaeal C-terminal stretch comprising the well conserved archaeal motif does not interfere with domain stability of yMBF1 (Figure 25) and moreover that the C-terminal part of eukaryotic MBF1 does not affect its function as shown previously the mutant Δ114-146 *Bm*MBF1 (TAKEMARU et al., 1997).

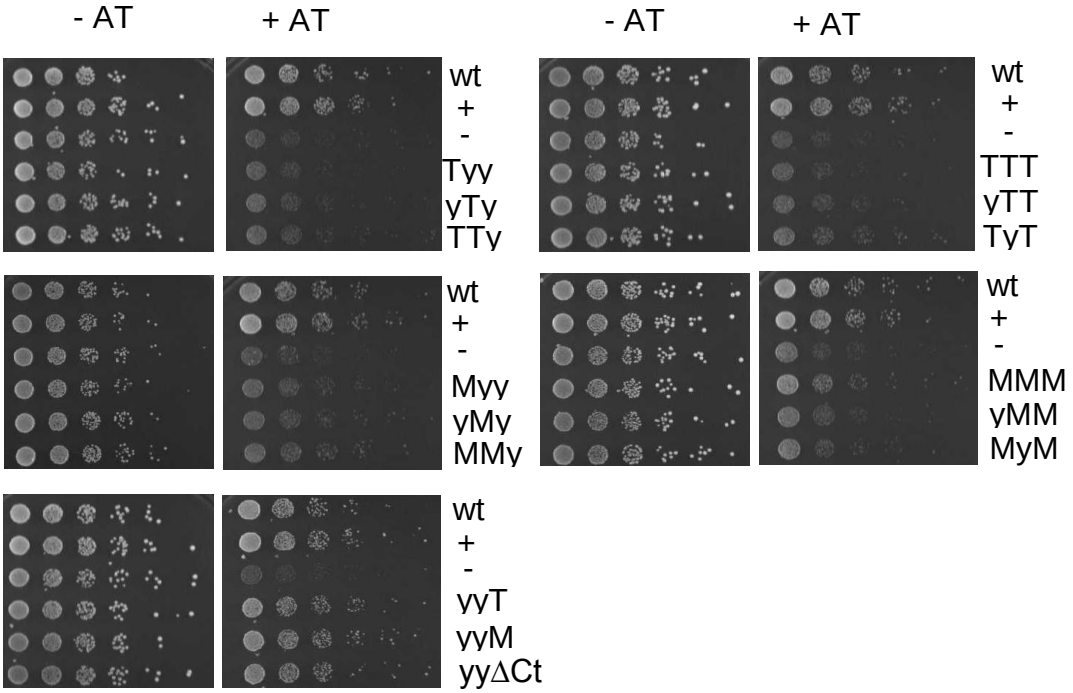


FIGURE 25. Complementation studies of yeast by MBF1 from yeast (y), aMBF1 (*T. tenax* (T) and *M. mazei* (M)), and the MBF1 chimera. *mbf1*Δ expressing MBF1 chimera containing the carboxyl-terminal domains of *T. tenax* or *M. mazei* MBF1 restored AT resistance. The yyΔCt mutant, which lacks the C-terminal stretch (residues 138 -151), was used as control. The expression of yeast MBF1 and chimera were assayed by growing the strains at YM agar plates in the presence of leucine and methionine (+Leu, +Met) as described above (Figure 23).

In summary, complementation studies with chimeric Archaea/yeast proteins indicate that neither the N-terminal, nor the core domain of Archaea is able to substitute for the respective eukaryotic domains. Only the archaeal and yeast C-terminal extension, which was shown here to be dispensable for yMBF1 function, can be successfully swapped. Since both, N-terminal and core domain of yeast are required for GCN4 interaction (TAKEMARU et al., 1998) and due to the absence of bZIP proteins in Archaea it might be predicted that archaeal MBF1 is not able to bind to the yeast activator GCN4.

The absence of bZIP-proteins like the transcriptional activator GCN4 in Archaea has been reported previously (KONING et al., 2009). In accordance with previous analysis, PSI-BLAST searches with yGCN4 (281 amino acids) in all available archaeal genomes revealed no homologues to bZIP-like regulator, GCN4. The most significant hits were 2-phospho-L-lactate transferase from *Methanoculleus marisnigri* JR1 (genbank accession: ref|YP_001048049.1| and gb|ABN58067.1|) and tryptophane synthase subunit beta from *Halobacterium* sp. NRC-1 (genbank accession: ref|NP_279407.1|) and from *Halobacterium salinarum* R1 (genbank accession: ref|YP_001688495.1|) with E-values of 0.02 and 0.86, respectively. Therefore, it is questionable, if the respective sites for GCN4 interactions are present in aMBF1 either in the divergent N-terminal domain or in the conserved HTH domain.

4. DISCUSSION

Although the third phylogenetic kingdom, the Archaea, are Prokaryotes, the processes involved in transferring genetic information, such as transcription, translation and DNA repair are more similar to their eukaryotic than to their bacterial counterparts (REEVE et al., 1997). The mechanisms and regulation of gene expression in Archaea have been studied in recent years, however, the knowledge is still rather limited compared to advances in transcription in Eukaryotes and Bacteria (BARTLETT, 2005; GEIDUSCHEK & OUHAMMOUCH, 2005).

4.1. BIOINFORMATIC ANALYSIS OF CRENARCHAEAL TFBS.

In the complete *T. tenax g*enome sequence (SIEBERS et al., manuscript in preparation) one TBP homologue (*Ttx*-TBP) and four TFB homologues (*Ttx*-TFB1-4) were identified. Homology searches using BLASTP reflect a high degree of sequence similarity within the crenarchaeal TFB1-group, euryarchaeal TFBs and eukaryotic TFIIB, whereas TFB2 and TFB3 protein sequences are only conserved within the phylum Crenarchaeota. For TFB4 no homologies were identified in archaeal genomes. Therefore, TFB4 is seems to be restricted to *T. tenax*.

A phylogenetic tree including 46 crenarchaeal TFBs suggests three distinct clades within Crenarchaeota, termed TFB1, TFB2 and TFB3 (Figure 2). All crenarchaeal species with available genome sequence possess both TFB1 and TFB2 members but TFB3 is not widely distributed in this phylum. The phylogenetic tree reflects the specification event of Thermoproteaceae, Sulfolobaceae and Desulfurococcaceae within each clade. Interestingly, TFB3 is encoded by almost all Sulfolobales and Thermoproteales genomes but is less distributed in Desulfurococcales. Strikingly, some species of the Thermoproteales family encode additional homologues of TFB3, for example, two in *P. calidifontis* and three in *P. arsenaticum*. Taken together, this points towards an ancient gene duplication, which results in the presence of three conserved TFB paralogues in almost all crenarchaeal species.

Similar to TFIIB, archaeal TFBs possess a structurally complex, conserved N-terminal region composed by a Zn-ribbon and B-finger motif that is connected by a linker to a globular C-terminus. The Zn-ribbon interacts with the RNAP “dock” domain during RNAP recruitment and it is postulated to have a crucial role in promoter opening and promoter escape by RNAP. The specific function of the B-finger in transcription initiation is still not clear. The N-terminus of TFB is located close to the transcription start site in the PIC, as shown by photochemical cross-linking experiments (BARTLETT et al., 2004).

Previous work using a truncated version of TFB1 in *S. acidocaldarius* and *S. solfataricus* (i.e. Δ1-40 *Sac*-TFB1 and Δ110 – 309 *Sso*-TFB1) revealed that the N-terminal domain (NTD) of TFB is necessary for RNAP recruitment. Experiments such as, yeast two-hybrid screenings and *in vitro* biochemical analyses have demonstrated that subunit K (RpoK) from *S. solfataricus* RNAP interacts with the NTD of TFB1 and is ultimately responsible for the association of RNAP with basal transcription factors and DNA (MAGILL et al., 2001; HICKEY et al., 2002). RpoK exhibits similarity with the ω-subunit of the bacterial RNAP and RBP6, a subunit shared by all three eukaryotic nuclear RNAPs (MINAKHIN et al., 2001; BEST & OLSEN, 2001). Therefore, it appears that RpoK represents an evolutionarily ancient component of the transcription systems of all three domains of life (MAGILL et al, 2001).

The C-terminal two-thirds of TFB comprise a helix-turn-helix motif that mediates the sequence-specific recognition of the BRE during transcription initiation and which is suggested to determine the polarity of the PIC. It has been previously shown that the C-terminal domain (CTD) structure in both *P. woesei* TFB (PwTFBc) and human TFIIB (hTFIIBc) have nearly identical folds, resembling two cyclin A-like domains (KOSA et al., 1997). The two cyclin-fold domains have positively charged surfaces that make direct contact with both TATA binding protein (TBP) and DNA residues upstream and downstream of the BoxA/TATA-box sequence (LANGER et al., 1995; KOSA et al., 1997). Also, the helix-turn-helix domain in the second repeat of the C-terminus is essential for both binding the TFIIB recognition element (BRE) and orientation of the transcription complex on promoter DNA (LAGRANGE et al., 1998).

Domain and structure analyses revealed that *Ttx*-TFB2-4 show significant modifications (Figure 26). *Ttx*-TFB1, as a classical TFIIB, exhibits a zinc finger near the N-terminus, a B-finger motif and two cyclin domains, also named TFB repeats. In *Ttx*-TFB2, the first cyclin domain is missing. However, secondary structure analysis revealed five α-helices between the B-finger and the "second" TFB-repeat at the exact positions of the α-helices of the "first" TFB-repeat, suggesting a similar function. *Ttx*-TFB3 is much shorter than the other homologues, which is also reflected in the domain structure. In *Ttx*-TFB3, the B-finger is missing. Directly after the zinc-ribbon structure, two α-helices are identified, followed by the "second" TFB-repeat. *Ttx*-TFB4 does not comprise a B-finger and also, no TFB-repeats can be identified; however, secondary structure prediction reveals two times five α-helical structures. In summary, even though the domain structures of these four *Ttx*-TFB homologues differ apparently, the secondary structures are very similar, with the exception of *Ttx*-TFB3.

While apparently essential domains are missing in *Ttx*-TFB2-4, the question arises, whether these proteins have an active function in transcription.

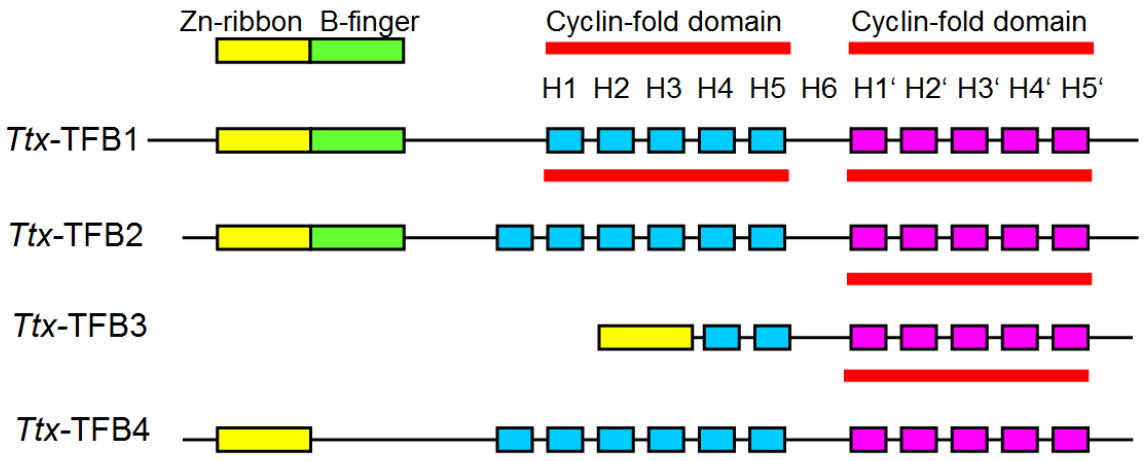


FIGURE 26. Schematic overview of the secondary and domain structure of *Ttx*-TFBs. Zn-ribbon: yellow box, B-finger: green box, Cyclin-fold domain: red bar. α-helices depicted as blue boxes for the first cyclin domain and pink boxes for the second cyclin domain.

Multiple sequence aligments, secondary and domain structure analyses of crenarchaeal TFBs (Figure 3 and 4) reflect that the Zn-ribbon motif is present in all crenarchaeal TFBs, except of TFB2 from *D. kamchatkensis*. TFB1s of Sulfolobales and Desulfurococcales exhibit a non-canonical Zn-ribbon motif (with the only exception of *I. hospitalis*), whereas the canonical Zn-ribbon motif is in all crenarchaeal TFB2s (except TFB2 from *M. sedula* which contains a non-canonical Zn-ribbon motif), TFB3s and TFB4 of *T. tenax*. The Zn – ribbon is followed by a conserved B-finger motif in TFB1s and TFB2s in Crenarchaeota, but is missing in TFB1 from *A. pernix*, TFB2 from *T. pendens* and *D. kamchatkensis* and also in all TFB3s and TFB4 of *T. tenax* (Figure 3, 4 and 26).

The CTD of all crenarchaeal TFBs are composed by α-helices, however, the two-cyclin fold domains are not detected in all members. The two TFB repeats are conserved in all TFB1s and in almost all TFB2s from Sulfolobales and Desulfurococcales (except TFB2 of *M. sedula* lacks the first one and TFB2 of *H. butylicus* lacks the second one) (Figure 3 and 4). The major modifications in domain structure of the CTD are found in TFB2 from Thermoproteales since both two-TFB repeats are missing in all members except TFB2 of *T. pendens* (both are present), and TFB2 of *T. tenax* and *P. islandicus*, which lack only the first cyclin domain (Figure 3 and 4). Interestingly, crenarchaeal TFB3s is the shortest homologue, because, it lacks the region corresponding to the first TFB repeat but the second cyclin fold is present in all of them.

The four metal binding amino acids in the Zn-ribbon of eukaryotic TFIIBs are either three cysteines (C) and one histidine (H), or four cysteines (CHEN et al., 2000). It has been suggested that the β-sheets in the TFIIB Zn-ribbon interact with RNAPII, while the binding of the active amino acids to zinc stabilizes the sheets (CHEN et al., 2000).

The second and third cysteines of the Zn-ribbon are well conserved in all crenarchaeal TFB1s (Figure 3). The first cysteine is replaced by serine in *S. solfataricus* and *S. acidocaldarius*, alanine in *S. tokodaii*, lysine in *D. kamtchatkensis* and *S. marinus*, aspartic acid in *M. sedula*, glycine in *H. butylicus* or glutamic acid in *A. pernix*. The last cysteine is substituted by threonine in Sulfolobales and Desulfurococcales TFB1 (except in *I. hospitalis*) (Figure 3). It was suggested previously that serine and threonine might also support the metal-binding capacity of the structure, because the oxygen atoms in these amino acids can interact with metal ions (QURESHI et al., 1995), but no experimental evidences are available so far. Taken together the present observations and the previous evidences from point mutation experiments in yeast and *S. solfataricus*, this indicates that both third and second cysteines might be essential for the structure of the Zn-ribbon motif of the TFBs.

Interestingly, the additional homologues present in crenarchaeal genomes, TFB2, TFB3 and TFB4, contain the two pairs of cysteine residues with the only exception of TFB2 from *D. kamchatkensis*, which lacks the Zn-ribbon (C18N, C21P, C37G and C40E) and *M. sedula* (the fourth cysteine is replaced by aspartic acid) (Figure 3).

It has been previously demonstrated that mutagenesis of residues C45 and L52 in the Zn-ribbon of yeast TFIIB impairs direct interaction with RNAP II (PARDEE et al., 1998). Site-directed mutagenesis illustrated that the analogous residues C31 and L38 in TFB of *Sulfolobus solfataricus* (C31A and L38P) are important for TFB – RpoK interaction, demonstrating the functional importance of this zinc-containing structure in Archaea (MAGILL et al., 2001). The analogous residues to L52 and L38 of the Zn-ribbon of yTFIIB and *Sso*-TFB1 is also conserved in other crenarchaeal TFBs or is replaced by the similar residues isoleucine and valine, except in *Tpe*-TFB1 (L55M), *Tpe*-TFB2 (L30F), *Sac*-TFB2 (L29D), *Dka*-TFB2 (L44Y) and *Mse*-TFB3 (L32A), which show unfavourable natural modifications, suggesting an impaired function of RpoK interaction in these homologues.

The mutation E46K located in the B-finger of *Ssa*-TFB1 affects the ability of RNAP to clear the promoter (BELL & JACKSON, 2000). The importance of Glu^{46} in the interaction with RNAP has also been confirmed for TFIIB of yeast and of human by mutational analysis (E62K) (PINTO et al., 1994; BANGUR et al., 1997; PARDEE et al., 1998) and E51K (CHO &

BURATOWSKI, 1999). The analogous residue to E46 of *Ssa*-TFB1 is well conserved in the B-finger of all crenarchaeal TFB1 and TFB2 but this residue is replaced by glycine in *Ape*-TFB1, by serine in *Tpe*-TFB2 and *Dka*-TFB2 and by cysteine in *Sso*-TFB2.
In addition, a helix-turn-helix (HTH) motif was demonstrated to establish a sequence-specific contact with the BRE in the *P. furiosus* TBP-TFB1-DNA co-crystal and the amino acids responsible for these contacts are Q268, V280 and R283. The respective residues to Q268 and R283 are conserved in all crenarchaeal TFB1s and TFB2s. R283 is additionally conserved in some homologues of TFB3 (i.e. *Pis*-TFB3, *Sac*-TFB3, *Pca*-TFB3b and the three homologues of *P. arsenaticum*). The respective residue of V280 is present in all crenarchaeal TFB1s and TFB2s from Sulfolobales and Desulfurococcales (except, *Hbu*-TFB2, the valine is replaced by alanine) but is not conserved in TFB2s of Thermoproteales and TFB3s (except, C*m*a-TFB2 and *Pca*-TFB3a).

Interestingly, gene context analyses in Archaea illustrate a conserved organization of *tfb1* and *tfb3* genes only in the family of Thermoproteaceae, contrasting with the high conservation of TFB1 homologues at protein level within the domain of Archaea. In contrast to TFB1, the genomic context of TFB2 is well conserved within the archaeal domain (Figure 5). In *T. tenax*, the *tfb4* gene is embedded between two genes non-related to informational processing mechanisms. Since TFB4 seems to be specific for *T. tenax* no conserved gene context is observed. Perhaps the *Ttx*-TFB4 performs different functions, recognizes different promoters or it represents a pseudogene. TFB1 and TFB4 proteins of *T. tenax* share 25% amino acid identity. Therefore, gene context analyses, sequence similarity and domain structure of *tfb4* suggest that *tfb4* is species specific homologue, which could be originated from a late gene duplication of the TFB1 homologue in *T. tenax*.

In summary, homology searches, gene context analysis, domain structure and phylogenetic analysis of crenarchaeal TFBs indicate that TFB1 is higher conserved within the group of Archaea and TFB2 and TFB3 only in Crenarchaeota suggesting that gene duplications occurred at an early stage of evolution. Taken together, these results suggest that TFB1 fulfils the function as housekeeping transcription factor similar to primary sigma factors in Bacteria, while the much more divergent TFB2 and TFB3 may play a regulatory role in crenarchaeal transcription, maybe as predicted previously (BALIGA et al., 2000) similar to the fast evolving alternative sigma factors in Bacteria.

The two conserved motifs, the Zn-ribbon and the B-finger present in the N-terminal region of the archaeal TFB and eukaryotic TFIIB are important for RNAP recruitment and transcription

start site selection (MICORESCU et al., 2008). The Zn-ribbon interacts with the RNAP "dock" domain during RNAP recruitment, but the specific function of the B-finger in the transcription mechanism is still not clear (COLANGELO et al., 2000). Because the HTH motif (cyclin domains) is necessary for interaction with the BRE and with TBP, and the N-terminal domain (NTD) is required for RNAP recruitment, it is tempting to speculate about the functions of those atypical TFBs, such as TFB2 from *T. tenax* which lack the first cyclin domain.

4.2 CHARACTERIZATION OF TFBs FROM *T. tenax*.

In order to study the function of multiple *Ttx*-TFBs, they were expressed in recombinant form. Thus, the ORFs encoding *Ttx*-TFB2, *Ttx*-TFB3 and *Ttx*-TFB4 were amplified by PCR from *T. tenax* genomic DNA and were cloned into the pET expression vectors. *Ttx*-TFB2 and *Ttx*-TFB3 were expressed as an N-terminally 6x-histidine-tagged protein and were purified to apparent homogeneity by heat treatment of the extract and chromatography on Ni^{2+}-Nitrilo-acetic acid-agarose.

In spite of the use of different expression vectors (pET15b, pET11c, pET24a and pQE-30) and different expression cells (Rosetta(DE3) and BL21(DE3)) no expression of *Ttx*-TFB4 was observed. In addition some modifications were introduced into the *tfb4* sequence via PCR mutagenesis in order to improve the expression of *Ttx*-TFB4 (see 3.2). However, all attempts were not successful. It is possible that *Ttx*-TFB4 expression can only be accomplished in an archaeal hyperthermophile or even in *T. tenax*, and not in a mesophilic bacterial host as *E. coli*.

The *tfb4* gene contains 24% of the codons for arginine (19), isoleucine (17) and leucine (27) (Figure 3) which are weakly translated in *E.coli*. From the 19 arginine residues in *Ttx*-TFB4, 11 are RIL codons AGA/AGG. The minor arginine $tRNA^{Arg}$ (AGG/AGA) has been shown to be a limiting factor in the bacterial expression of several mammalian genes (BRINKMANN et al., 1989), because the codons AGA and AGG are infrequently used in *E. coli*. Furthermore, the *tfb4* sequence contains a tandem AGGAGG triplet, which has been suggested to cause a substantial inhibition of gene expression independent of their localization in mRNA. This inhibitory effect has been attributed to a competition of the tandem AGGAGG codons with the natural Shine Dalgarno sequence (MAKRIDES, 1996). Therefore, it is not surprisingly that the expression of *Ttx*-TFB4 in *E. coli* has been not successful.

The many advantages of *E. coli* have ensured that it remains a valuable organism for the high-level production of recombinant proteins. However, in spite of the extensive knowledge on

the genetics and molecular biology of *E. coli*, not every gene can be expressed efficiently in this organism. This may be due to the unique and subtle structural features of the gene sequence, the stability and translational efficiency of mRNA, the ease of protein folding, degradation of the protein by host cell proteases, major differences in codon usage between the foreign gene and native *E. coli*, and the potential toxicity of the protein to the host. There are no exact rules that ensure the complete succes of heterologous expression in *E. coli* but general guidelines have emerged (MAKRIDES, 1996). For example, AUG is the preferred codon by two- to threefold, and GUG is only slightly better than UUG (RINGQUIST et al., 1992). The Shine-Dalgarno (SD) site, also known as ribosomal binding sequence (RBS) interacts with the complementary 3' end of 16S rRNA during translation initiation. The spacing between the SD site and the initiating AUG codon can vary from 5 to 13 nucleotides, and it influences the efficiency of translational initiation (GOLD, 1988).

In *E. coli*, there is a preference for the UAA stop codon because the major release factors, RF-1 and RF-2, terminate translation at this stop codon (GRENTZMANN et al., 1994). It has been demonstrated that termination efficiencies varied significantly depending on both the stop codon and the fourth nucleotide, ranging from 80% (UAAU) to 7% (UGAC). This indicates that the identity of the nucleotide immediately following the stop codon strongly influences the efficiency of translational termination in *E. coli* (POOLE et al., 1995). Therefore, UAAU is the most efficient translational termination sequence in *E. coli*. The design of expression vectors frequently includes the insertion of all three stop codons to prevent possible ribosome skipping (MAKRIDES, 1996).

Ttx-TBP, *Ttx*-TFB1 and also *Ttx*-TFB2 were cloned, expressed and purified previously by F. BLOMBACH (2005). Briefly, the proteins were expressed as native-length recombinant proteins, lacking any tags. *Ttx*-TBP and *Ttx*-TFB2 were purified to apparent homogeneity by heat treatment of the extract, followed by chromatography on a Q-Sepharose anion exchange column and heparin sepharose column, respectively. *Ttx*-TFB1 aggregates in form of inclusion bodies. Therefore, the membrane fraction containing *Ttx*-TFB1 was dissolved on denaturating conditions and *Ttx*-TFB1 was refolded and purified to apparent homogeneity by heat treatment followed by heparin sepharose chromatography.

To characterize the purified recombinant TFBs from *T. tenax*, the binding capacity to DNA was studied. Electrophoretic mobility shift assays (EMSAs) with *Ttx*-TFB1, *Ttx*-TFB2 and *Ttx*-TFB3 and *fba-pfp* promoter in presence and absence of *Ttx*-TBP, were performed.

As shown previously for homologues of TFB1 and TBP from other archaeal species and also eukaryotic counterparts, both factors together generated stable protein-DNA complex

formation in *T. tenax* as well, whereas *Ttx*-TFB1 or *Ttx*-TBP alone exhibits essentially no binding to DNA (Figure 9 and 12A). These results indicate that the presence of TFB1 and TBP is essential for stable DNA binding, as reported previously for eukaryotic TFIIBs and for TFB1 homologues of other archaeal species (SOPPA et al., 1999; LITTLEFIELD et al., 1999; BELL & JACKSON, 2000; HAUSNER & THOMM, 2001). Surprisingly, *Ttx*-TFB2 was able to form a stable complex in absence of *Ttx*-TBP and no supershift was observed in the presence of *Ttx*-TBP (Figure 11 and 12B). These observations of *Ttx*-TFB1 and *Ttx*-TFB2 binding were confirmed also by S. KONING in the laboratory of Prof. Dr. BETTINA SIEBERS for other DNA promoter regions (i.e. *tfb1*, *gar1-tfb2*, *pps* and *orf1155-lrp*) regulating genes encoding proteins with diverse function.

Strinkingly, *Ttx*-TBP provokes a negative effect on the *Ttx*-TFB2-DNA complex formation. EMSAs revealed that the banding pattern generated in presence of *Ttx*-TBP is identical only at low *Ttx*-TFB2 amounts (1.2 and 2 ng) (Figure 11, lane 2, 3 and 7, 8). The formation of corresponding high molecular size DNA-protein complexes in presence of *Ttx*-TBP required the addition of higher amounts of *Ttx*-TFB2 (Figure 11, lane 4 and 9). This negative effect of *Ttx*-TBP on the *Ttx*-TFB2 complex formation was confirmed in several independent experiments (Figure 11, 12 and 13).

One possibility is that *Ttx*-TFB2 might interact with *Ttx*-TBP, which recruits part of the free-*Ttx*-TFB2 and therefore, less *Ttx*-TFB2 is available to interact with DNA. However, no ternary complex *Ttx*-TFB2/*Ttx*-TBP/DNA was observed by EMSAs. Notably, previous evidences indicate that the two cyclin domains of archaeal TFBs make direct contact with TBP. It is tempting to speculate that due to the structure modification of *Ttx*-TFB2, which lacks the first cyclin fold domain, the complex *Ttx*-TFB2/*Ttx*-TBP might interact with specific AT-rich target DNA sequences and not with a wide-range of BRE/TATA boxes as the complex *Ttx*-TFB1/*Ttx*-TBP does (i.e. BRE/TATA box of *fba-pfp*). This effect could function as a kind of mechanism for transcription regulation. A second possibility is that *Ttx*-TBP and *Ttx*-TFB2 might compete for the same binding site at the DNA (i.e. AT-rich sites). As previously reported *Ttx*-TBP-DNA complexes are unstable and therefore, only *Ttx*-TFB2-DNA complexes are detected in EMSAs.

The specificity of the DNA-binding of *Ttx*-TFB1 and *Ttx*-TFB2 to the *fba-pfp* promoter fragment was studied by EMSAs using wild type short-length fragments and the respective mutated sequences as DNA probe and, by competition experiments. The *fba-pfp* promoter region was selected as DNA probe to study specificity of DNA binding, because this operon has been studied in detail previously. The transcription start site of the *fba-pfp* operon has

been determined and the operon was shown to be differentially expressed (i.e. up-regulated under heterotrophic growth conditions) (SIEBERS et al., 2004). The predicted *fba-pfp* promoter is located between positions -38 and -23 and was identified by comparison with the consensus crenarchaeal promoter, which was predicted by SLUPSKA and co-workers (2001) based on computational analysis of the regions from -50 to +50 nt of 10 unrelated protein-encoding genes from the hyperthermophilic Crenarchaeon *P. aerophilum* (SLUPSKA et al., 2001).

Interestingly, *Ttx*-TFB1 and *Ttx*-TFB2 bind to a short length DNA probe of 60 bp comprising position +50 to -10 of the *fba-pfp* promoter fragment (Figure 12). As observed for the large length DNA probe (265 bp) *Ttx*-TFB1 binds to the 60 bp-length *fba-pfp* fragment in a TBP dependent manner and *Ttx*-TFB2 binds stable in presence or absence of TBP (Figure 12). Additionally also *Ttx*-TFB3 shows *Ttx*-TBP independent binding as observed for *Ttx*-TFB2 (Figure 12C). Competition experiments using the unlabeled 60 bp-length *fba-pfp* fragment illustrate that increasing amounts of specific competitor DNA reduce the TFB2-DNA complex formation and the competition was significantly enhanced in the presence of *Ttx*-TBP (Figure 13).

In addition, two mutated sequences of the 265 bp fragment encompassing the *fba-pfp* promoter region were used as DNA probes for EMSAs to assess the specificity of DNA binding. In *fba-pfpP*-265bp the mutations are located between positions -55 to +15 of the *fba-pfp* promoter region and in *fba-pfpB/T*-265bp between positions -35 to -22. Surprisingly, in the presence or absence of *Ttx*-TBP and increasing amounts of *Ttx*-TFB1 or *Ttx*-TFB2, binding was observed to both mutated sequences. Notably, the *Ttx*-TFB1/*Ttx*-TBP binding appeared to be unstable to both mutated DNA probes, since only a smear was observed during electrophoresis (Figure 15A). Thus, although the BRE/TATA box of the *fba-pfp* promoter is mutated, the *Ttx*-TFB1/*Ttx*-TBP is able to bind DNA, however unstable, suggesting unspecific binding or binding to alternative sequences. In contrast to TFB1, binding of *Ttx*-TFB2 to the mutated sequences, *fba-pfpB/T*-265bp and *fba-pfpP*-265bp in presence or absence of *Ttx*-TBP is not affected (Figure 15B and C). Notably the binding is stable as observed previously using wild-type fragments. These results, using DNA fragments containing mutations on the *fba-pfp* promoter region, therefore, revealed that the mutations on BRE/TATA box does not affect *Ttx*-TFB2 binding to the *fba-pfp* DNA probe and suggest that *Ttx*-TFB2 binds either unspecific or stable to sequences outside of *fba-pfp* core promoter. To our knowledge this is the first report, where a TFB homologue binds stable to DNA independent of TBP.

To identify DNA binding sites of *Ttx*-TFB1 and *Ttx*-TFB2 Exo III footprinting assays were performed. In this study no binding of *Ttx*-TBP alone to DNA was detected by Exo III footprinting assays (Figure 17) confirming the aforementioned observations by EMSAs. Binding to the predicted BRE/TATA box of *fba-pfp* fragment was only observed in the presence of *Ttx*-TFB1 and *Ttx*-TBP as shown by the Exo III stop signals at positions -16 and +1 which were not detected in presence of *Ttx*-TFB1 alone. These results indicate that *Ttx*-TFB1 in the presence of *Ttx*-TBP binds stable to the core promoter as a typical archaeal TFB. Homologues of TFB or TBP from other archaeal species and Eukaryotes yielded essentially similar results: no binding to the promoter was observed for single GTFs and only in the presence of both factors a DNase I footprint was generated; for instance, *Sac*-TFB and *Sac*-TBP cover the region from -43 to -14 of the SSV T6 promoter (BELL & JACKSON, et al., 2000).

Strinkingly, no binding of *Ttx*-TFB2 to the *fba-pfp* BRE/TATA box in presence or absence of *Ttx*-TBP was found in this study by Exo III footprinting analysis (Figure 17). Therefore, *Ttx*-TFB2 seems not to be able to bind to the *fba-pfp* core promoter in the presence of *Ttx*-TBP as *Ttx*-TFB1 does. In Eukaryotes as in Archaea, the N-terminal region of TFIIB (TFB) plays a key role in the recruitment of RNAP to promoters but not in the interaction with TBP. For example, a truncated version of TFB from *S. acidocaldarius* lacking the entire N-terminal region is able to form a ternary complex with TBP on the SSV T6 promoter with comparable efficiency as the wild-type *Sac*-TFB. The DNase I footprinting indicated that the protection pattern induced by the mutant *Ssa*-TFB was indistinguishable from that induced by the wild-type *Ssa*-TFB. However, the mutant *Ssa*-TFB was unable to mediate RNAP recruitment (BELL & JACKSON, et al., 2000).

It was also observed within this work that the presence of *Ttx*-TFB1 or *Ttx*-TFB2 alone revealed several additional stops of digestion with Exo III in the upstream region of the BRE/TATA box on the *fba-pfp* coding strand. The major predominant boundaries are at positions -76 and -58. The derived region for protein binding comprises a AT-rich binding sequence ([-101] 5'-TCAATAAAATCATAAAA-3' [-84]) with high similarity to the archaeal consensus promoter (ANAAAA(NNN)CTTTTAAA) and to the *fba-pfp* core promoter ([-38] 5'-ACAAAAAGATATTAAAA-3' [-21]; only 5 bp difference in 17 bp). No difference between Exo III footprinting patterns of *Ttx*-TFB2/DNA complex in absence and in presence of *Ttx*-TBP was detected confirming the aforementioned observations of the EMSA.

It is important to mention that a direct repeat of 6 bases is detected in the predicted AT-rich sequence of the *fba-pfp* operon. Interestingly, the analysis of the different promoter DNA

fragments used as probes for EMSAs illustrates that the promoter region of the *pps* gene also contains an AT-rich sequence around the position (-80) with an inverted palindromic repeat of 6 bases. These predicted consensus upstream sequences of *fba-pfp* and *pps* seems to be too far away from the transcription start sites to serve as promoter however, for a final decision further experimental studies are required.

It is noteworthy that *Ttx*-TFB1 binding to the mutated *fba-pfp* promoter is unstable as revealed by EMSAs using the *fba-pfpP*-265bp (mutations located from position -55 to +15) and the *fba-pfpB*/T-265bp (mutations located from position -38 to -21) probe. The mutations target specifically the BRE/TATA box (*fba-pfpB*/T-265bp) and a larger region (from 17 bp upstream to 36 bp dowstream of the BRE/TATA box). The AT-rich region (-101 to -84) in both mutated sequences is not modified. *Ttx*-TFB2 showed stable binding to the mutated sequences and the binding is not affected by the presence of TBP. Binding of *Ttx*-TFB2 to AT-rich sites is also supported by the fact that *Ttx*-TBP might compete with *Ttx*-TFB2 for the same binding site (AT-rich sites) as observed in EMSAs (Figure 11, 12 and 13). Therefore, it is tempting to speculate that the -101/-84 sequence functions as an upstream activating sequence (UAS) where TFB2 binds. These results suggest that *Ttx*-TFB2 does not function as basal transcription factor (i.e. typical TFIIB), but rather as transcriptional activator; similar to the function of *Sso*-TFB3 (PAYTUBI & WHITE, 2009).

In summary, *Ttx*-TFB1 and *Ttx*-TFB2 share 33% sequence similarity. *Ttx*-TFB1 behaves as a typical TFIIB(TFB), which exhibits a conserved structure and a TBP dependent DNA binding, whereas *Ttx*-TFB2 shows significant structural modifications. The N-terminal domain of *Ttx*-TFB2 is well-conserved encompassing the canonical Zn-ribbon motif and B-finger, but harbors a modified C-terminal cyclin-fold comprising the core domain lacking the first cyclin-fold. The natural modified *Ttx*-TFB2 and its particular TBP independent DNA binding found in this study therefore, seems to support the structural prediction findings that the C- terminal core domain of TFB is essential for the interaction with TBP and not the N-terminal domain, which has been reported to be important for RNAP recruitment. Thus, Exo III footprinting analysis, EMSAs and structure analyses suggest that *Ttx*-TFB2 does not function as basal transcription factor (i.e. typical TFIIB), but rather as transcriptional activator.

In several archaeal genomes different regulators have been identified, which their regulatory mechanisms are based on UAS. For example, different regulatory binding sites have been identified for members of the Lrp family and this is in agreement with most of the data obtained from footprinting experiments with a variety of Lrp-like proteins, where stretches of

hundred bp or more are typically protected against cleavage. High-affinity binding sequences for Lrp homologues do contain sequence elements with dyad symmetry, and it was demonstrated that specific Lrp regulators require specific bases at certain positions (CUI et al., 1995; OUHAMMOUCH & GEIDUSCHEK, 2001) (Figure 28).

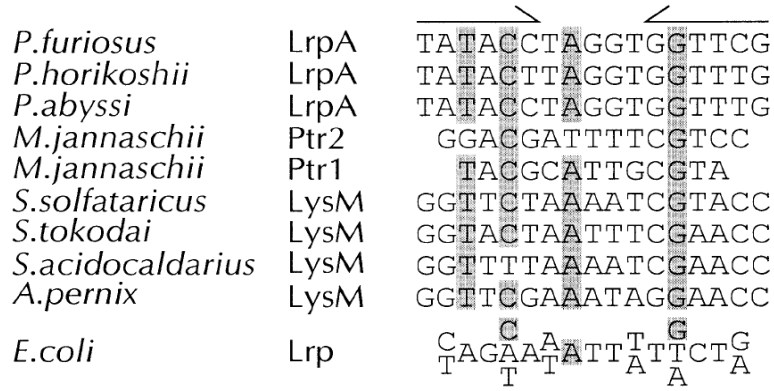


FIGURE 27. Lrp binding site. The binding sites of Lrp-like proteins from *M.jannaschii* Ptr1, Ptr2 and *E. coli* Lrp were determined by SELEX (TUERK and GOLD, 1990; CUI et al., 1995; OUHAMMOUCH & GEIDUSCHEK, 2001); binding sites for *P. furiosus* LrpA and *S. solfataricus* LysM were determined using footprinting analysis (BRINKMAN *et al*., 2000; 2002); binding sites for their paralogues were determined by comparison of the respective promoter sequences. Horizontal arrows indicate (partial) inverted repeat elements. Taken from Brinkman et al., 2003.

These present studies in *T. tenax* are the first, which address the function of TFB2 in the phylum Crenarchaeota. They underline the importance of the TATA box and the AT-rich region in gene expression and the involvement of multiple transcription factors, such as *Ttx*-TFB2 and other regulatory proteins in the activation and modulation of transcription. In general, the mechanism involving GTF multiplicity might have evolved to handle the great environmental dynamics experienced in extreme environments. However, for a more detailed understanding of the suggested activation mechanism and regulation further genetic and biochemical analyses are required, including the development of an *in vitro* transcription system. This will be an important requisite for understanding the function of multiple TFBs in the crenarchaeal domain, i.e. *T. tenax*.

4.3 MULTIPROTEIN BRIDGING FACTOR 1.

Due to the eukaryotic-like nature of the basal transcription machinery in Archaea, it might be predicted that archaeal regulators are eukaryotic- type regulators, however, transcriptional regulation in Archaea involves a mosaic of DNA-binding proteins frequently of the bacterial type, modulating a eukaryotic-type core transcription apparatus. Most of the known regulators

in Archaea are negative regulators with only a few exceptions, such as Ptr1 and Ptr2 from *M. janashii*, Mth from *M. thermoautotrophicum*, MDR-1 from *A. fulgidus*, Lrs14 from *S. solfataricus* and LrpA from *P. woesei* (see introduction).

Interestingly, multiprotein bridging factor 1 (MBF1) is absent in Bacteria, but is well-conserved in Archaea and Eukaryotes (ARAVIND & KOONIN, 1999). MBF1 has been identified in all archaeal genomes with available genome sequence information. Diverse biological functions of MBF1 have been reported in Eukaryotes (LI et al., 1994; TAKEMARU et al., 1998; SMITH et al., 1998; KABE et al., 1999; TSUDA et al., 2004), but its function in Archaea remains still unclear. In Eukaryotes MBF1 is a transcriptional co-activator, which connects transcription activators with TBP (TAKEMARU et al., 1997, 1998). It has been suggested previously that MBF1 is a required basal transcription factor, because of its degree of conservation in Eukaryotes and the presence of homologues in Archaea (ARAVIND & KOONIN, 1999). The function of co-activators from Archaea has not been studied so far and the biological role of aMBF1 as multiprotein bridging factor has not been documented.

The interacting activator partners of MBF1 belong either to the steroid/nuclear hormone receptor family (e.g. FTZ-F1 in insects (TAKEMARU et al., 1997) or are leucine zipper (bZIP)-type transcriptional activators, such as GCN4 in yeast (TAKEMARU et al., 1998) (see introduction). In yeast, MBF1 is essential for GCN4-dependent activation of *HIS3*, encoding an imidazole glycerol-phosphate dehydratase that is responsible for the third step in histidine biosynthesis. Without activation, a basic level of constitutive expression of this gene still allows growth in yeast. Disruption of the *mbf1* gene generates a phenotype that is sensitive to 3-aminotriazole, an inhibitor of the enzyme imidazole-3-phosphate dehydratase (HIS3) (TAKEMARU et al., 1998).

In order to unravel the role of MBF1 in Archaea, *MBF1* from the hyperthermophile *T. tenax* (T*MBF1*) and the mesophile *Methanosarcina mazei* (M*MBF1*) as well as Archaea-yeast chimeric *MBF1* variants were analyzed for their ability to complement the function of yeast *MBF1*. The mesophilic *M. mazei* MBF1 was included in this study in order to exclude temperature-dependent effects, which might be caused by expression in a mesophilic host (e.g. incorrect folding, missing activity). The two archaeal orthologs of MBF1 exhibit 36% sequence identity. As shown in Figure 23A, complementation of yeast *mbf1*Δ (y*mbf1*Δ) by either the mesophilic or hyperthermophilic archaeal MBF1, was not successful.

The soluble expression of TMBF1 in yeast deletion strain was confirmed by western blot analysis (Figure 23B). In general, genes encoding different proteins from extremophiles have been cloned in mesophilic hosts, such as, *E. coli*, *Bacillus subtilis* and yeast (CIARAMELLA et

al., 1995). For example, the RNA polymerase subunit P from *P. furiosus* functionally replaced the eukaryotic polymerase subunit Rbp12 in a respective yeast mutant (REICH et al., 2009). More recently TFB1, TFB3 and all the subunits of RNAP from *S. solfataricus* P2 were efficiently expressed in yeast and used for yeast-two-hybrid experiments (PAYTUBI & WHITE, 2009).

Thus, it appears most likely that TMBF1 is formed in a correct folded and active state, which is further supported by the fact that also the mesophilic aMBF1 of *M. mazei* was not able to complement. The expression of each of the three paralogues of MBF1 from the plant *Arabidopsis thaliana* was able to restore MBF1 function in a yeast deletion strain (*mbf1*Δ) (TSUDA et al., 2004). Moreover, the defect of a yeast *mbf1* disruptant was rescued by expression of a cDNA encoding silkworm or human MBF1 (TAKEMARU and HIROSE, unpublished observation (TAKEMARU et al., 1997). Therefore, aMBF1, in contrast to eukaryal MBF1 from human, insects (TAKEMARU et al., 1997) or plants (TSUDA et al., 2004), is not able to complement and to restore eukaryal MBF1 wild-type function (AT resistance) in yeast by substituting for eukaryal MBF1 during GCN4-mediated transcription activation.

Domain swap experiments revealed that the divergent archaeal N-terminal domain and the conserved core domain of MBF1, harboring the conserved HTH domain, are not functional in yeast (Figure 25). This was surprising, since the HTH domain of MBF1 is a cro-HTH type domain and has been identified previously as the only highly conserved, classical HTH domain that is vertically inherited in all Archaea and Eukaryotes (ARAVIND & KOONIN, 1999). Furthermore, the results obtained in this study indicate that the archaeal C-terminal extension, containing the well conserved archaeal motif "[TS]-[LIVMF]-G-[DEN]" does not interfere with yMBF1 function and revealed that the C-terminal extension of yeast MBF1 is not required for MBF1 function (yyΔCt-MBF1) (Figure 25). To our knowledge, this is the first experimental study to gain insights into the physiological function of MBF1 in Archaea (manuscript in preparation).

The regions of MBF1 required for protein – protein interaction with TBP and the sequence specific activator have been identified using bacterially expressed MBF1 truncated protein in insects and in yeast (TAKEMARU et al., 1997; 1998). Deletion analyses of *B. mori* MBF1 and yMBF1 demonstrated that the central region (residues 35–113 and residues 41–119, respectively) is essential for the binding of TBP and the transcriptional activator, FTZ-F1 in *B. mori* (TAKEMARU et al., 1997) and GCN4 (TAKEMARU et al., 1998) in yeast. The C-terminal half of this region (residues 65–132 and residues 71–138) is markedly conserved among various species and seems to contribute, mainly to the interaction, because deletion of

the N-terminus (residue 1 – 66) and the C-terminus from residue 114 still retains significant activity in *Bm*MBF1. NMR studies revealed the presence of four amphipathic helices in this highly conserved region (residues 63–132) in *Bm*MBF1, that are bundled into a compact form (OZAKI et al., 1999).

In Eukaryotes NMR data from hMBF1 and *Bm*MBF1 proposed that the HTH-core domain is responsible for maintaining domain stability (OZAKI et al., 1999; MISHIMA et al., 1999). Binding of TBP in hMBF1 occurs via the amphipathic helices in the core HTH domain (MISHIMA et al., 1999) and in yeast an aspartate residue at position 112 (Asp 112), located in helix III, was identified as important binding site for TBP (TAKEMARU et al., 1998).

Interestingly, this region is also highly conserved in Archaea (Figure 20) and was predicted as the only highly conserved cro-HTH domain vertically inherited from Archaea to Eukarya (ARAVIND & KOONIN, 1999). The N-terminal half of this essential region (residues 35–62 in *Bm*MBF1 and residues 41–68 in yMBF1) exhibits less prominent conservation and in Archaea high level of basic residues (arginine or lysine) are located in the linker (Figure 21). In *B. mori* it was predicted that this region might play a functional role rather than an architectural one (TAKEMARU et al., 1997), because of its assumed flexible structure.

The N-terminal domain of MBF1 from *Bm*MBF1, yMBF1 and hMBF1 is not required for binding to TBP. In insects and in yeast, binding of TBP to a truncated version of MBF1 lacking its N-terminal amino acids (residues 1–34 in *Bm*MBF1 (TAKEMARU et al., 1997) and residues 1–40 in yMBF1 (TAKEMARU et al., 1998)) remained unaffected, but the binding of the activator, GCN4 or FTZ-F1, was slightly reduced. Therefore, both, the N-terminal region and the core domain are indispensable for the binding to the activator, whereas, the core domain is essential for binding to TBP. The observation that the N-terminal domain of MBF1 is not required for binding to TBP in human, yeast and insects (TAKEMARU et al., 1997; 1998; LI et al., 1994) suggests that at least aMBF1 and yTBP interaction is possible, despite the divergent N-terminal region with a conserved Zn-ribbon motif.

Based on a bioinformatic and mutational approach in yeast Liu and co-workers (LIU et al., 2007) suggested the co-evolution of MBF1 with TBP in Eukaryotes and Archaea. Aspartic acid (position 112) in yMBF1 and glutamine (position 68) in yTBP were identified as important residues required for MBF1:TBP interaction during GCN4-dependent transcriptional activation (TAKEMARU et al., 1998; LIU et al., 2007). In aMBF1 they identified lysine, arginine, serine, or asparagine as interacting residues corresponding to yMBF1 D112 based on the multiple alignment of 21 full-length MBF1 protein sequences (LIU et al., 2007). The *in vivo* mutational approach for yMBF1 revealed that combinations that occur in nature

(e.g. the mutant yMBF1-D112R (WT, TBP-68Q); Archaea) showed no effect, whereas unnatural combinations like yMBF1-D112K (WT, TBP-68Q) were sensitive to AT, indicating that the interactions between yMBF1 and yTBP were disrupted in the latter mutants (LIU et al., 2007). The amino acids corresponding to position 112 in TMBF1 and MMBF1, are arginine (position 102) and lysine (position 108), respectively (Figure 20), suggesting that TMBF1 but not MMBF1 is able to interact with yTBP. However, as shown in this study TMBF1 is not able to complement the yeast *mbf1*Δ strain in spite of the presence of the appropriate amino acid for the interaction with yTBP.

Beside aMBF1-yTBP interaction also the binding of aMBF1 to the activator GCN4, which depends on both, the N- terminal and HTH- domain, might be affected. In yeast the binding of GCN4 to MBF1 was not altered by a D112A mutation (TBP binding), but it was slightly reduced by the N-terminal deletion of yMBF1 (*MBF1*ΔNT). Both the D112A and *MBF1*ΔNT mutants are sensitive to AT as shown by TAKEMARU and co-workers (1998).

Interestingly, it has been reported previously that Archaea lack bZIP-proteins like the transcriptional activator GCN4 (KONING et al., 2009). This observation was also confirmed in the present study. Therefore, it might be possible that the respective sites for GCN4 interactions are not conserved in aMBF1, neither in the divergent N-terminal domain nor in the conserved HTH domain.

The conservation of MBF1 among all organisms, in which TBP is used as general transcription factor, suggests that MBF1 has a fundamentally important function also in Archaea. However, the absence of bZIP-like proteins in the archaeal domain, the presence of a Zn-ribbon and positive- charged linker in the aMBF1 and the complementation experiments using Archaea- yeast chimeric protein as presented within this study, suggests a different function –may be as single transcriptional regulator– of MBF1 in Archaea. A possible role of MBF1 in transcription or translation has been proposed previously (KONING et al. 2009), however, further investigations are required to elucidate the role of MBF1 in Archaea.

5. SUMMARY

Archaeal transcription is generally regarded as simpler model of eukaryotic transcription (BALIGA et al., 2000). Relatively little is known about the interaction of gene-specific regulators and the basal transcription apparatus and about how transcriptional regulation in Archaea helps to confer fitness across a broad range of environments, including hostile ones. Multiple forms of the general transcription factors (GTFs), TBP and TFB, are commonly found in Archaea (LANGER et al., 1995). It has been suggested previously that multiple TBPs and TFBs might function similar to bacterial sigma-factors, which regulate transcription in response to environmental changes (PAYTUBI & WHITE, 2009). In the group of Crenarchaeota, only TFB1 from *S. solfataricus* and *S. acidocaldarius* (BELL & JACKSON, 2000) and recently, TFB3 from *S. solfataricus* (PAYTUBI & WHITE, 2009) have been studied biochemically. However, the role of TFB2 in Crenarchaeota is still unknown and no *tfb* knockout mutants are available for any of the crenarchaeal TFBs.

The hyperthermophilic crenarchaeon *T. tenax g*enome encodes one TBP and four TFBs. Bioinformatic analyses revealed that *Ttx*-TFB1 harbors a Zn-ribbon motif, B-finger and two cyclin domains similar to other TFB homologues in Archaea, however, *Ttx*-TFB2-4 possess significant modifications in secondary structure. In order to study the function of *Ttx*-GTFs, recombinant proteins were expressed in *E. coli*. EMSAs were performed using the *fba-pfp* promoter and revealed that *Ttx*-TFB1, as classical TFB, binds to DNA stable in a TBP-dependent manner whereas *Ttx*-TFB2 and *Ttx*-TFB3 shows stable DNA binding independent of TBP. To our knowledge, this is the first report of a TFB homologue binding DNA independent of TBP.

Exo III footprinting analyses using the *fba-pfp* promoter (encoding fructose-1,6-bisphophate aldolase and PP_i-dependent phosphofructokinase in *T. tenax* (SIEBERS et al., 2004)) illustrate that *Ttx*-TFB1 binds to the core promoter (from position -38 to -21), which is stabilized by *Ttx*-TBP. However, for *Ttx*-TFB2, no binding to the BRE/TATA box of the *fba-pfp* promoter region and no influence of *Ttx*-TBP is observed. Furthermore, *Ttx*-TFB1 and *Ttx*-TFB2 bind in addition to a sequence (from position -101 to -84) with high similarity to BRE/TATA box of the archaeal consensus promoter (SLUPSKA et al., 2001). Mutations of the *fba-pfp* promoter region affected the stability of the binding of *Ttx*-TFB1/*Ttx*-TBP complex to DNA suggesting that the complex binds unstable to the region -101/-84. The stable binding of *Ttx*-TFB2 to the region -101/-84 suggest that this region might function similar to an upstream activation sequence (UAS) for the *fba-pfp* promoter of *T. tenax*. Therefore, these results suggest that *Ttx*-

TFB2 does not function as basal transcription factor (i.e. typical TFIIB), but rather as transcriptional activator; similar to the function of *Sso*-TFB3 (PAYTUBI & WHITE, 2009).

Multiprotein bridging factor 1 (MBF1) is a transcriptional co-activator that bridges a sequence-specific activator and TBP in Eukaryotes. MBF1 is absent in Bacteria, but is well-conserved in Eukaryotes and Archaea and harbors a C-terminal Cro-like Helix Turn Helix (HTH) domain, which is the only highly conserved, classical HTH domain that is vertically inherited in all Eukaryotes and Archaea. Phylogenetic analyses revealed a common distribution of MBF1 in all Archaea with available genome sequence. The main structural difference between archaeal MBF1 (aMBF1) and eukaryotic MBF1 is the presence of a Zn-ribbon motif in aMBF1. The function of co-activators from Archaea has been not studied so far and the biological role of aMBF1 as multiprotein bridging factor has not been documented.

In yeast the MBF1 mediated activation of histidine synthesis by the bZIP-like activator GCN4 has been well studied (TAKEMARU et al., 1998). To study the function and therefore the evolutionary conservation of MBF1 and its single domains, complementation studies in yeast (*mbf1Δ*) as well as domain swap experiments between aMBF1 and yMBF1 were performed. In contrast to previous reports for eukaryal MBF1 (i.e. *Arabidopsis thaliana*, insect and human), the two archaeal MBF1 orthologs, TMBF1 from the hyperthermophile *T. tenax* and MMBF1 from the mesophile *M. mazei* were not functional for complementation of a yeast mutant lacking MBF1 (*mbf1Δ*). From 12 chimeric proteins representing different combinations of the N-terminal, core domain, and the C-terminal extension from yeast and aMBF1, only the chimeric MBF1 comprising the yeast N-terminal and the core domain fused to the archaeal C-terminal part was able to restore full wild-type activity of MBF1. The yMBF1 mutant lacking the C-terminal part was also able to restore MBF1 activity, suggesting that this part is not important for MBF1 function.

The absence of bZIP-like proteins in the archaeal domain, the presence of a Zn-ribbon in the divergent N-terminal domain of aMBF1 and the complementation experiments using Archaea- yeast chimeric proteins presented here suggests no obvious conservation in biological function of MBF1 within Archaea and Eukaryotes. It is tempting to speculate that aMBF1 might act as a single regulator, which directly interacts with DNA for transcriptional regulation via its Zn-ribbon motif. To our knowledge, this is the first experimental study to gain insight about the physiological function of MBF1 in Archaea.

6. LITERATURE

ALBRIGHT, S.R. AND TJIAN, R. (2000). TAFs revisited: More data reveal new twists and confirm old ideas. Gene 242: 1–13.

ALTING-MEES, M.A., SHORT, J.M. (1989). pBluescript II: gene mapping vectors, Nucleic Acids Res. 17: 9494.

ALTING-MEES, M.A., SORGE, J.A., SHORT, J.M. (1992). pBluescriptII: multifunctional cloning and mapping vectors, Meth. Enzymol. 216: 483-495.

ALTSCHUL, S.F., MADDEN, T.L., SCHAFFER, A.A., ZHANG, J., ZHANG, Z., MILLER, W., LIPMAN, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of generation of protein database search programs. Nucleic Acids Res. 25: 3389-3402.

ARAVIND, L., KOONIN, E.V.(1999). DNA-binding proteins and evolution of transcription regulation in the archaea. Nucleic Acids Res. 27 (23): 4658-4670.

BALIGA, N.S., GOO, Y.A., NG, W.V., HOOD, L., DANIELS, CH.J., DASSARMA, SH. (2000). Is gene expression in Halobacterium NRC-1 regulated by multiple TFB and TFB transcription factors?. Mol Microbiol. 36: 1184-1185.

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. Mol. Cell. Biol. 17: 6784-6793.

BARANGER, A.M., PALMER, C.R., HAMM, M.K., GIEBLER, H.A., BRAUWEILER, A., NYBORG, J.K., SCHEPARTZ, A. (1995). Mechanism of DNA-binding enhancement by the human T-cell leukaemia virus transactivator Tax. Nature. 376: 606-608.

BARTLETT, M. S. (2005). Determinants of transcription initiation by archaeal RNA polymerase. Curr. Opin. Microbiol. 8: 677–684.

BARTLETT, M.S., THOMM, M., GEIDUSCHEK, E.P. (2004). Topography of the euryarchaeal transcription initiation complex. J. Biol. Chem. 279: 5894–5903.

BELL, S.D., JAXEL, CH., NADAL, M., KOSA, P.F., JACKSON, S.P. (1998). Temperature, template topology, and factor requirements of archaeal transcription.Proc. Nalt. Acad. Sci. USA. 95: 15218-15222.

BELL, S.D., CAIRNS, S.S., ROBSON, R.L., JACKSON, S.P. (1999). Transcriptional regulation of an archaeal operon in vivo and in vitro. Mol. Cell. 4: 971-982.

BELL, S.D., JACKSON, S.P. (2000). The role of transcription factor B in transcription initiation and promoter clearance in the archaeon *Sulfolobus acidocaldaricus*. J. Biol. Chem. 275: 12934-12940.

BELL, S. D., BRINKMAN, A. B., VAN DER OOST, J., AND JACKSON, S. P. (2001). The archaeal TFIIEα homologue facilitates transcription initiation by enhancing TATA-box recognition. EMBO Rep. 2: 133–138

BELL, S.D., JACKSON, S.P. (2001). Mechanism and regulation of transcription in archaea. Curr. Opin. Microbiol. 4: 208–213.

BERK, A.J. (2000). TBP-like factors come into focus. Cell. 103: 5–8.

BEST, A.A., OLSEN, G.J. (2001). Similar subunit architecture of archaeal and eukaryotic RNA polymerases. FEMS Microbiol. Lett. 195: 85–90.

BLOMBACH, F. (2005). Establishment of an *in vitro* transcription system for the hyperthermophilic crenarchaeum *Thermoproteus tenax*. Bachelor Thesis. University Duisburg-Essen.

BOYD, J., OZA. M.N., MURPHY, J.R. (1990). Molecular cloning and DNA sequence analysis of diphtheria *tox* iron-dependent regulatory elements (*dtxH)* from *Corynebacterium diphtheriae*. Proc. Natl. Acad. Sci. USA 87: 5968–5972.

BRADFORD, M.M. (1976). A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem. 72: 248-254.

BREATHNACH, R., CHAMBON, P. (1981). Organisation and expression of split genes coding for proteins. Annu. Rev. Biochem. 50: 49-83.

BRENDEL, C., GELMAN, L., AUWERX, J. (2002). Multiprotein bridging factor-1 (MBF-1) is a cofactor for nuclear receptors that regulate lipid metabolism. Mol. Endocrinol. 16: 1367-1377.

BRINKMAN, A.B., DAHLKE, I., TUININGA, J.E., LAMMERS, T., DUMAY, V., DE HEUS, E., et al. (2000). An Lrp-like transcriptional regulator from the archaeon *Pyrococcus furiosus* is negatively autoregulated. J Biol Chem. 275: 38160–38169.

BRINKMAN, A.B., BELL, S.D., LEBBINK, R.J., DE VOS, W.M., VAN DER OOST, J. (2002) The *Sulfolobus solfataricus* Lrp-like protein LysM regulates lysine biosynthesis in response to lysine availability. J Biol Chem. 277: 29537–29549.

BRINKMAN, A.B., ETTEMA, T.J.G., DE VOS, W.M., VAN DER OOST, J. (2003). Lrp-like transcription regulators. Mol Microbiol. 48 (2): 287–294.

BRINKMANN, U., MATTES, R. E., BUCKEL, P. (1989). High-level expression of recombinant genes in *Escherichia coli* is dependent on the availability of the *dnaY* gene product. Gene 85:109–114.

BROCHIER-ARMANET, C., BOUSSAU, B., GRIBALDO, S., FORTERRE, P. (2008). Mesophilic crenarchaeota: proposal for a third archaeal phylum, the thaumarchaeota. Nat. Rev. Microbiol. 6: 245–252.

BURATOWSKI, S. (1997). Multiple TATA-Binding Factors come back into style. Cell 91: 13-15.

BURKE, T.W. AND KADONAGA, J.T. 1996. Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. Genes & Dev. 10: 711–724.

BURLEY, S.K. AND ROEDER, R.G. 1996. Biochemistry and structural biology of transcription factor IID (TFIID). Annu. Rev. Biochem. 65: 769–799.

BUSHNELL, D.A., WESTOVER, K.D., DAVIS, R.E., KORNBERG, R.D. (2004). Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 angstroms. Science 303: 983–988.

BUTLER, J.E.F., KADONAGA, J.T. (2002). The RNA polymerase II core promoter: a key component in the regulation of gene expression. Genes Dev. 16: 2583-2592.

CARSTENS, C.P., WAESCHE, A. (1999). Codon bias-adjusted BL21 derivatives for protein expression. Strategies Newsletter 12.

CHEN, H-T., LEGAULT, P., GLUSHKA, J., OMICHINSKI, J.G., AND SCOTT, R.A. (2000). Structure of a (Cys3His) Zn-ribbon, a ubiquitous motif in archaeal and eukaryotic transcription. Prot. Sci. 9: 1743–1752.

CHO, E.J., BURATOWSKI, J. (1999). Evidence that Transcription Factor IIB is required for a post-assembly step in transcription initiation. J. Biol. Chem. 274: 25807-25813.

CHOMCZYNSKI, P., MACKEY, K., DREWS, R., WILFINGER, W. (1993). A reagent for the rapid isolation of genomic DNA. Biotechniques. 22: 550-553.

CIARAMELLA, M., CANNIO, R., MORACCI, M., PISANI, F.M., ROSSI, M. (1995). Molecular biology of extremophiles. World J. Microbiol. Biotechnol. 11:71–84.

COKER, J.A., DASSARMA, S. (2007). Genetic and transcriptomic analysis of transcription factor genes in the model halophilic archaeon: coordinate action of TBPD and TfbA. BMC Gen. 8: 61-73.

COLANGELO, CH.M., LEWIS, L.M., COSPER, N.J., SCOTT, R.A. (2000). Structural evidence for a common zinc binding domain in archaeal and eukaryotic transcription factor IIB proteins. JBIC. 5: 276-283.

CUI, Y., WANG, Q., STORMO, G.D., CALVO, J.M. (1995) A consensus sequence for binding of Lrp to DNA. J Bacteriol. 177: 4872–4880.

DAHLKE, I., THOMM, M. (2002). A *Pyrococcus* homolog of the leucineresponsive regulatory protein, LprA, inhibits transcription by abrogating RNA polymerase recruitment. Nucleic Acids Res. 30: 701–710.

DANTONEL, J.CH., WURTZ, J.M., POCH, O., MORAS, D., TORA, L. (1999). The TBP-like factor: alternative transcription factor in Metazoa?. Trends Biochem Sci. 24: 335-339.

DeLONG, E.F., DAVID M.K. (2005). Genomic perspectives in microbial oceanography. Nature 437: 336-342.

DeLONG, E.F (1998) Everything in moderation: archaea as 'non-extremophiles. Curr Opin Genet. 6:649–54.

DRAGONI, I., MARIOTTI, M., CONSALEZ, G.G., SORIA, M.R., MAIER, J.A. (1998). EDF-1, a novel gene product down-regulated in human endothelial cell differentiation. J. Biol. Chem. 273: 31119-31124.

FAIGER, H., IVANCHENKO, M., COHEN, I., HARAN, T.E. (2006). TBP flanking sequences: asymmetry of binding, long-range effects and consensus sequences. Nucleic Acids Res. 34:104-119.

FELSENSTEIN, J. (1985) Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783-791.

FIORE, A.D., FLORENTINO, G., VITALE, R.M, RONCA, R., AMADEO, P., PEDONE, C., BARTOLUCCI, S., DE SIMONA, G. (2009) Structural Analysis of BldR from Sulfolobus solfataricus Provides Insights into the Molecular Basis of Transcriptional Activation in Archaea by MarR Family Proteins. J. Mol. Biol. 388: 559–569.

FREY, G., THOMM, M., BRIIDIGAM, B., GOHL, H.P., HAUSNER, W. (1990). An archaebacterial cell-free transcription system. The expression of tRNA genes from *Methanococcus uannielii* is mediated by a transcription factor. Nucleic Acids Res. 18: 1361-1367.

FRÖLS, S., GORDON, P.M., PANLILIO, M.A., DUGGIN, I.G., BELL, S.D., SENSEN, C.W., SCHLEPER, C. (2007). Response of the hyperthermophilic archaeon Sulfolobus solfataricus to UV damage. J Bacteriol 189: 8708–8718.

GEIDUSCHEK, E. P. & OUHAMMOUCH, M. (2005). Archaeal transcription and its regulators. Mol. Microbiol. 56: 1397–1407.

GRABOWSKI, B., KELMAN, Z. (2003). Archeal DNA replication: eukaryotic proteins in a bacterial context. Annu. Rev. Microbiol. 57: 487-516

GRENTZMANN, G., BRECHEMIER-BAEY, D., HEURGUE, V., MORA, L., BUCKINGHAM, R. H. **(**1994). Localization and characterization of the gene encoding release factor RF3 in *Escherichia coli*. Proc. Natl. Acad. Sci. USA 91:5848–5852.

GODOY, A.V., ZANETTI, M.E., SAN SEGUNDO, B., CASALONGUE, C. (2001). Identification of a putative *Solanum tuberosum* transcriptional co-activator up-regulated in potato tubers by *Fusarium solani* f. sp. eumartii infection and wounding. Physiologia Plantarum 112: 217-222.

GOEDE, B., NAJI, S., KAMPEN, O., ILG, K., THOMM, M. (2006). Protein-Protein Interactions in the Archaeal Transcriptional Machinery. Binding studies of isolated RNA polymerase subunits and transcription factors. J Biol Chem. 281(41): 30581–30592.

GOHL, H.P., GROENDAHL, B., THOMM, M. (1995). Promoter recognition in archaea is mediated by transcription factors: identification of transcription factor aTFB from *Methanococcus thermolithotrophicus* as archaeal TATA-binding protein. Nucleic Acids Res. 23: 3837-3841.

GOLD, L. **(**1988). Posttranscriptional regulatory mechanisms in *Escherichia coli*. Annu. Rev. Biochem. 57:199–233.

GRAHAM, D.E., OVERBEEK, R., OLSEN, G.J., WOESE, C.R. (2000). An archaeal genomic signature. Proc. Natl. Acad. Sci. 97: 3304–3308.

HANAHAN, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol. 166: 557-580.

HANZELKA, B. L., DARCY, T. J., AND REEVE, J. N. (2001). TFE, an Archaeal Transcription Factor in *Methanobacterium thermoautotrophicum* Related to Eukaryotic Transcription Factor TFIIEα. J. Bacteriol. 183: 1813–1818.

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*. J. Mol. Biol. 222: 495-508.

HAUSNER, W., THOMM, M. (2001). Events during initiation of archaeal transcription: open complex formation and DNA-protein interaction. J Bacteriol. 183: 3025-3031.

HENDRICK, J.L., WILSON, P.G., EDELMAN, II., SANDBAKEN, M.G., URSIC, D., CULBERTSON, M.R. (2001). Yeast frameshift suppressor mutations in the genes coding for transcription factor MBF1p and ribosomal protein S3: evidence for autoregulation of S3 synthesis. Genetics 157: 1141–1158.

HICKEY, A.J., DE MACARIO, E.C., MACARIO, A.J.L. (2002). Transcription in the Archaea: Basal Factors, Regulation, and Stress-Gene Expression. Crit. Rev. Biochem. Mol. Biol. 37(6): 537–599.

HÜDEPOHL, U., REITER, W.D., ZILLIG, W. (1990). *In vitro* transcription of two rRNA genes of the archaebacterium *Sulfolobus* sp. B 12 indicates a factor requirement for specific initiation. Proc. Natl. Acad. Sci. USA. 87: 5851-5855.

HIRATA, A., KLEIN, B. J. & MURAKAMI, K. S. (2008). The X-ray crystal structure of RNA polymerase from Archaea. Nature 451: 851–854.

HUFFMAN, J.L., BRENNAN, R.G. (2002). Prokaryotic transcription regulators: more than just the helix-turn-helix motif. Curr. Opin. Struct. Biol. 12: 98–106.

JINDRA, M., GAZIOVA, I., UHLIROVA, M., OKABE, M., HIROMI, Y., HIROSE, S. (2004). Co-activator MBF1 preserves the redox-dependent AP-1 activity during oxidative stress in *Drosophila*. EMBO J. 23: 3538-3547.

KABE, Y., GOTO, M., SHIMA, D., IMAI, T., WADA, T., MOROHASHI, K-I., SHIRAKAWA, M., HIROSE, S., HANDA, H. (1999). The role of human MBF1 as a transcriptional co-activator. J. Biol. Chem. 274 (48): 34196-34202.

KAUFMANN, J. AND SMALE, S.T. 1994. Direct recognition of initiator elements by a component of the transcription factor IID complex. Genes & Dev. 8: 821–829.

KEELING, P.J., CHARLEBOIS, R.L., DOOLITTLE, W.F. (1995) Archaebacterial genomes: eubacterial form and eukaryotic content. Curr Opin Genet Dev 4:816–822

KESSLER, A., SEZONOV, G., INAKI GUIJARRO, J., DESNOUES, N., ROSE, T., DELEPIERRE, M. et al. (2006). A novel archaeal regulatory protein, Sta1, activates transcription from viral promoters. Nucleic Acids Res. 34: 4837–4845.

KLENK, H-P., DOOLITTLE, W.F. (1994). Archaea and eukaryotes versus bacteria? Curr Biol 4:920–922.

KONING, B., BLOMBACH, F., WU, H., BROUNS, S.J.J., VAN DER OOST, J. (2009). Role of multiprotein bridging factor 1 in the archaea: bridging the domains?. Bioch. Soc. Trans. 37 (part 1): 52-57.

KOSA, P.F., GHOSH, G., DEDECKER, B.S., SIGLER, P.B. (1997). DNA recovered in 2.1-angstrom crystal structure of an archaeal preinitiation complex: TATA-box-binding protein/transcription factor (II)B core/TATA-box. Proc. Natl. Acad. Sci. USA 94: 6042–6047.

KUSSER, A. G., BERTERO, M. G., NAJI, S., BECKER, T., THOMM, M., BECKMANN, R. & CRAMER, P. (2008). Structure of an archaeal RNA polymerase. J. Mol. Biol. 376: 303–307.

KYRPIDES, N., OUZOUNIS, C. (1995). The eubacterial transcriptional activator Lrp is present in the Archaeon *Pyrococcus furiosus*. Trends Biochem Sci 20:140–141.

LAEMMLI, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685.

LANGER, D., HAIN, J., THURIAUX, P., ZILLIG, W. (1995). Transcription in Archaea-similarity to that in Eukarya. Proc Nat Acad Sci USA 92: 5768-5772.

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 Genes Dev. 12: 34-44.

LEPP, P.W., BRINIG, M.M., OUVERNEY, C.C., PALM, K., ARMITAGE, G.C., RELMAN, D.A. (2004). Methanogenic archaea and human periodontal disease. Proc Natl Acad Sci USA. 101(16):6176-6181.

LI, F.Q., UEDA, H., HIROSE, S. (1994). Mediators of activation of fushi tarazu gene transcription by BmFTZ-F1. Mol. Cell. Biol. 14: 3013-3021.

LITTLEFIELD, O., KORKHIN, Y., SIGLER, P.B. (1999). The structural basis for the oriented assembly of a TBP/TFB/promoter complex. Proc. Natl. Acad. Sci. USA 96:13668–13673.

LIU, Q., NAKASHIMA-KAMIMURA, N., IKEO, K., HIROSE, S., GOJOBORI, T. (2007). Compensatory change of interacting amino acids in the coevolution of transcriptional co-activator MBF1 and TATA-box binding protein TBP. Mol. Biol. Evol. 24: 1458–1463.

MACKEY, K., WILLIAMS, P., SEIM, S., CHOMCZYNSKI, P. (1996). The use of DNAzol for the rapid isolation of genomic DNA from whole blood. Biomed Prod Suppl. 263: 13-15.

MAGILL, C.P., JACKSON, S.P., AND BELL, S.D. (2001). Identification of a conserved archaeal RNA polymerase subunit contacted by the basal transcription factor TFB. J. Biol. Chem. 276: 46693– 46696.

MAKRIDES, S.C. (1996). Strategies for Achieving High-Level Expression of Genes in *Escherichia coli*. Microbiol. Rev. 60(3): 512-538.

MALDONADO, E. (1999). Transcriptional functions of a new mammalian TATA-binding protein-related factor. J Biol Chem. 274: 12963-12966.

MARIOTTI, M., DE BENEDICTIS, L., AVON, E., MAIER, J.A. (2000). Interaction between endothelial differentiation-related factor-1 and calmodulin *in vitro* and *in vivo*. J Biol Chem. 275: 24047-24051.

MARKOWITZ, VM., CHEN, I-M.A., PALANIAPPAN, K., CHU, K., SZETO, E., GRECHKIN, Y., RATNER, A., ANDERSON, I., LYKIDIS, A., MAVROMATIS, K., IVANOVA, N.N., KYRPIDES, N.C. (2009). The integrated microbial genomes system: an expanding comparative analysis resource. Nucleic Acids Res. doi:10.1093/nar/gkp887: 1–9.

MARTINEZ, E., CHIANG, C.-M., GE, H., AND ROEDER, R.G. 1994. TATA-binding protein-associated factor(s) in TFIID function through the initiator to direct basal transcription from a TATA-less class II promoter. EMBO J. 13: 3115–3126.

MATSUSHITA, Y., MIYAKAWA, O., DEGUCHI, M., NISHIGUCHI, M., NYUNOYA, H. (2002) Cloning of a tobacco cDNA coding for a putative transcriptional co-activator MBF1 that interacts with the tomato mosaic virus movement protein. J. Exp. Botany. 53: 1531-1532.

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*. J Bacteriol. 190: 157-167.

MINAKHIN, L., BHAGAT, S., BRUNNING, A., CAMPBELL, E. A., DARST, S. A., EBRIGHT, R. H., AND SEVERINOV, K. (2001). Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. *Proc. Natl. Acad. Sci. U. S. A.* 98**:** 892–897.

MIOTTO, B., STRUHL, K. (2006). Differential gene regulation by selective association of transcriptional co-activators and bZIP DNA-binding domains. Mol. Cel. Biol. 26(16): 5969-5982.

MISHIMA, M., OZAKI, J., IKEGAMI, T., KABE, Y., GOTO, M., UEDA, H., HIROSE, S., HANDA, H., SHIRAKAWA, M. (1999). Resonance assignments, secondary structure and 15N relaxation data of the human transcriptional co-activator hMBF1 (57–148). J. Biomol. NMR 14: 373–376.

NAPOLI, A., VAN DER OOST, J., SENSEN, C.W., CHARLEBOIS, R.L., ROSSI, M., CIARAMELLA, M. (1999). An Lrp-like protein of the hyperthermophilic archaeon *Sulfolobus solfataricus* which binds to its own promoter. J. Bacteriol. 181: 1474–1480.

OELGESCHLÄGER, T., CHIANG, C.-M., AND ROEDER, R.G. 1996. Topology and reorganization of a human TFIID-promoter complex. Nature 382: 735–738.

OLSEN, G.J., WOESE, C.R. (1996). Lessons from an Archaeal genome: what are we learning from *Methanococcus jannaschii*? Trends Genet. 12: 377–379.

OUHAMMOUCH, M., GEIDUSCHEK, E.P. (2001). A thermostable platform for transcriptional regulation: the DNA-binding properties of two Lrp homologues from the hyperthermophilic archaeon *Methanococcus jannaschii.* EMBO J. 20: 146–156.

OUHAMMOUCH, M., GEIDUSCHEK, E.P. (2005). An expanding family of archaeal transcriptional activators. PNAS. 102(43): 15423–15428.

OUZOUNIS, C., KYRPIDES, N., SANDER, C. (1995) Novel protein families in archaea genomes. Nucleic Acids Res. 23:565–570.

OUZOUNIS, CH.A., KYRPIDES, N.C. (1996). Parallel Origins of the Nucleosome Core and Eukaryotic Transcription from Archaea. J Mol Evol. 42: 234–239.

OZAKI, J., TAKEMARU, K., IKEGAMI, T., MISHIMA, M., UEDA, H., HIROSE, S., KABE, Y., HANDA, H., SHIRAKAWA, M. (1999). Identification of the core domain and the secondary structure of the transcriptional co-activator MBF1. Genes Cells 4: 415–424.

PAN, H.Q., WANG, Y.P., CHISSOE, S.L., BODENTEICH, A., WANG, Z., IYER, K., CLIFTON, S.W., CRABTREE, J.S., ROE, B.A. (1994). The complete nucleotide sequences of the SacBII Kan domain of the P1 pAD10-SacBII cloning vector and three cosmid cloning vectors: pTCF, svPHEP, and LAWRIST16. Gent Anal Tech Appl. 11: 181-186.

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. J. Biol. Chem. 273: 17859-17864.

PAYTUBI, S., WHITE, M. (2009). The crenarchaeal DNA damage-inducible transcription factor B paralogue TFB3 is a general activator of transcription. Mol Microbiol. Doi: 10.1111/j.1365-2958.2009.06737.x.

PERINI, G., WAGNER, S., GREEN, M.R. (1995). Recognition of bZIP proteins by the human T-cell leukaemia virus transactivator Tax. Nature 376: 602-605.

PINTO, I., WU, W-H., NA, J.G., HAMPSEY, M. (1994). Characterization of *sua7* mutations defines a domain of TFIIB involved in transcription start site selection in yeast. J Biol Chem. 269: 30569-30573.

POOLE, E. S., BROWN, C. M., TATE, W. P. **(**1995). The identity of the base following the stop codon determines the efficiency of *in vivo* translational termination in *Escherichia coli*. EMBO J. 14:151–158.

QURESHI S.A., KHOO, B., BAUMANN, P., JACKSON, S.P. (1995). Molecular cloning of the transcription factor TFIIB homolog from *Sulfolobus shibatae*. Proc. Natl. Acad. Sci. USA. 92: 6077– 6081.

QURESHI, S.A, BELL, S.D., JACKSON, S.P. (1997). Factor requirements for transcription in the archaeon *Sulfolobus shibatae*. EMBO J. 16: 2927-2936.

REEVE, J.N., SANDMAN, K., DANIELS, C.J. (1997). Archaeal histones, nucleosomes, and transcription initiation. Cell 89 (7): 999–1002.

REICH, CH., ZELLER, M., MILKEREIT, P., HAUSNER, W., CRAMER, P., TSCHOCHNER, H., THOMM, M. (2009). The archaeal RNA polymerase subunit P and the eukaryotic polymerase subunit Rpb12 are interchangeable in vivo and in vitro. Mol. Microbiol. 71(4): 989-1002.

RINGQUIST, S., SHINEDLING, S. BARRICK, D. GREEN, L., BINKLEY, J. STORMO, G. D., GOLD, L. **(**1992). Translation initiation in *Escherichia coli*: sequences within the ribosome-binding site. Mol. Microbiol. 6:1219–1229.

ROWLANDS, T., BAUMANN, P., JACKSON, S.P. (1994). The TATA-binding protein: A general transcription factor in Eukaryotes and Archaebacteria. Science 264: 1326-1329.

SAITOU, N., NEI, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4:406-425.

SAMBROOK, J., FRITSCH, E.F., MANIATIS, T. (1989). Molecular Cloning- A Laboratoty Manual, 2nd Edition.

SANGER, F., NICKLEN, S., COULSON, A.R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci USA. 74: 5463-5467.

SANTANGELO, T.J., CUNONOVA, L., JAMES, C.L., REEVE, J.N. (2007). TFB1 or TFB2 sufficient for Thermococcus kodakarensis viability and for basal transcription in Vitro. J Mol Biol. 367: 344-357.

SCHIRALDI, CH., GIULIANO, M.T., DE ROSA, M. (2002). Perspectives on biotechnological applications of archaea. Archaea 1: 75–86.

SCHNABEL, R., THOMM, M., GERARDY-SCHAHN, R., ZILLIG, W., STETTER, K. O., HUET, J. (1983). Structural homology between different archaebacterial DNA-dependent RNA polymerases analysed by immunological comparison of their components. EMBO J. 2: 751-755.

SIEBERS, B., SCHÖNHEIT, P. (2005). Unusual pathways and enzymes of central carbohydrate metabolism in Archaea. Curr. Opin. Microbiol. 8: 695–705.

SIEBERS, B., TIADEN, B., MICHALKE, K., DOERR, C., AHMED, H., ZAPARTY, M., GORDON, P., SENSEN, C.W., ZIBAT, A., KLENK, H.P., SCHUSTER, S.C., HENSEL, R. (2004). Reconstruction of the central carbohydrate metabolism of *Thermoproteus tenax* by use of genomic and biochemical data. J. Bacteriol. 186: 2179-2194.

SLUPSKA, M.M., KING, A.G., FITZ-GIBBON, S., BESEMER, J., BORODOVSKY, M., MILLER, J.M. (2001). Leaderless Transcripts of the Crenarchaeal Hyperthermophile *Pyrobaculum aerophilum*. J. Mol. Biol. 309: 347-360.

SMITH, M.L., JOHANSON, R.A., ROGERS, K.E., COLEMAN, P.D., SLEMMON, J.R. (1998). Identification of a neuronal calmodulin-binding peptide, CAP-19, containing an IQ motif. Brain Res. Mol. Brain Res. 1998, 62: 12-24.

SOPPA, J. (1999). Transcription initiation in Archaea: facts, factors and future aspects Mol. Microbiol. 31(5): 1295-1305.

SÖDING, J. (2005) Protein homology detection by HMM-HMM comparison. Bioinformatics 21: 951-960.

STUDIER, F.W., MOFFATT, B.A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J Mol Biol. 189: 113-130.

SUZUKI, N., BAJAD, S., SHUMAN, J., SHULAEV, V., MITTLER, R. (2008). The transcriptional co-activator MBF1c is a key regulator of thermotolerance in *Arabidopsis thaliana*. J. Biol. Chem. 283(14): 9269-9275.

TAKEMARU, K.I., UEDA, H., HIROSE, S. (1997) Multiprotein bridging factor 1 (MBF1) is an evolutionary conserved transcriptional co-activator that connects a regulatory factor and TATA element-binding protein. Proc. Natl. Acad. Sci. U.S.A. 94: 7251-7256.

TAKEMARU, K-I., HARASHIMA, S., UEDA, H., HIROSE, S. (1998). Yeast Co-activator MBF1 mediates GCN4-dependent transcriptional activation. Mol. Cel. Biol. 18 (9): 4971-4976.

TAMURA, K., DUDLEY, J., NEI, M., KUMAR, S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. 10.1093/molbev/msm092.

TJADEN B, PLAGENS A, DÖRR C, SIEBERS B, HENSEL R. (2006). Phosphoenolpyruvate synthetase and pyruvate, phosphate dikinase of *Thermoproteus tenax*: key pieces in the puzzle of archaeal carbohydrate metabolism. Mol Microbiol. 60(2):287-298.

THOMM, M. (1996). Archaeal transcription factors and their role in transcription initiation. FEMS Microbiol Rev. 18: 159-171.

THOMPSON, J.D., GIBSON, T.J., PLEWNIAK, F., JEANMOUGIN, F., HIGGINS, D.G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25: 4876-4882.

TSUDA, K., TSUJI, T., HIROSE, S., YAMAZAKI, K-I. (2004). Three Arabidopsis MBF1 homologues with distinct expression profiles play roles as transcriptional co-activators. Plant Cell Physiol. 45(2): 225-231.

TUERK, C., GOLD, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science 249: 505–510.

VENTER, J.C., et.al. (2004). Environmental Genome Shotgun Sequencing of the Sargasso Sea. Science 304:66-74.

VERRIJZER, C.P. 2001. Transcription factor IID—Not so basal after all. Science 293: 2010–2011.

WANG, Q., CALVO, J.M. (1993). Lrp, a major regulatory protein in *Escherichia coli*, bends DNA and can organize the assembly of a higher-order nucleoprotein structure. EMBO J. 12: 2495–2501.

WRAY, G.A., HAHN, M.W., ABOUHEIF, E., BALHOFF, J.P., PIZER, M.., ROCKMAN, M.V., ROMANO, L.A. (2003). The evolution of transcriptional Regulation in Eukaryotes. Mol. Biol. Evol. 20(9):1377–1419.

WERNER, F., WEINZIERL, R. O. (2005). Direct modulation of RNA polymerase core functions by basal transcription factors. Mol. Cell Biol. 25:8344–8355.

WETTACH, J., GOHL, H.P., TSCHOCHNER, H., THOMM, M. (I995). Functional interaction of yeast and human TATA binding proteins with an archaeal RNA-polymerase and promoter. Proc. Natl. Acad. Sci. USA. 92: 472-476.

WHITMAN, W.B., PFEIFER, F., BLUM, P., KLEIN, A. (1999). What Archaea Have to Tell Biologists. Genetics. 152: 1245–1248.

WOESE, C.R., FOX, G.E. (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc. Natl. Acad. Sci. USA. 74: 5088–5090.

ZANETTI, M.E., BLANCO, F.A., DALEO, G.R., CASALONGUÉ, C.A. (2003). Phosphorylation of a member of the MBF1 transcriptional co-activator family, StMBF1, is stimulated in potato cell suspensions upon fungal elicitor challenge. J. Exp. Botany. 54(383): 623-632.

[a]ZAPARTY, M., TJADEN, B., HENSEL, R., SIEBERS, B. (2008). The central carbohydrate metabolism of the hyperthermophilic crenarchaeote *Thermoproteus tenax*: pathways and insights into their regulation. Arch. Microbiol. 190: 231–245.

[b]ZAPARTY, M., ZAIGLER, A., STAMME, C., SOPPA, J., HENSEL, R., SIEBERS, B. (2008). DNA microarray analysis of the central carbohydrate metabolism: glycolytic/gluconeogenic carbon switch in the hyperthermophilic Crenarchaeum *Thermoproteus tenax*. J Bacteriol. 190(6): 2231–2238.

ZEGZOUTI, H., JONES, B., FRASSE, P., MARTY, C., MAITRE, B., LATCHÉ, A., PECH, J.C., BOUZAYEN, M. (1999). Ethylene-regulated gene expression in tomato fruit: characterization of novel ethylene-responsive and ripening-related genes isolated by differential display. Plant J. 18: 589-600.

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 Zn-ribbon. Nat. Struct. Biol. 3: 122-124.

ZILLIG, W., STETTER, K.O., SCHÄFER, W., JANEKOVIC, D., WUNDERL, S., HOLZ, I., PALM, P. (1981). Termoproteales: a novel type of extremely thermoacidophilic anaerobic archaebacteria isolated from Icelandic solfatares. Zentbl Bakteriol Hyg 1 Abt Org C2: 205-227.

ZILLIG, W., STETTER, K.O., JANEKOVIC, D. (1979). DNA-dependent RNA polymerase from the archaebacterium *Sulfolobus acidocaldarius*. Eur J Bioch. 96: 597-604.

7. APPENDIX

7.1 LIST OF ABREVIATIONS

aa	Amino acid
A. bidest	Aqua bidestillata = two times distilled water
AT	Aminotriazole
BLAST	Basic Local alignment search tool
bp	bases pair
BV	Bootstrap value
bZIP	basic region/leucine zipper
CCM	Central carbohydrate metabolism
CE	Crude extract
CTD	C-terminal domain
DNA	Desoxyribonucleic acid
DPE	Downstream core promoter element
e.g.	for example
EDF1	Endothelial differentiation factor 1
EDTA	Ethylene diamine tetraacetic acid
EMSA	Electrophoretic mobility shift assay
et al.	et alii = and the others
Exo III	Exonuclease III
fba	fructose bisphosphate aldolase gene
GTF	general transcription factor
HP	heat precipitation
HTH	Helix turn helix
Inr	Initiator element
IPTG	Isopropyl-β-D-thiogalactopyranoside
KDa	Kilodalton
L	Liter
LB	Lysogeny broth (popularly known as Luria Bertani, Lennox broth)
Lrp	Leucine responsive regulatory protein
M	Molar (mol/L)
mA	Milliampere
MBF1	Multiprotein bridging factor 1
min	Minute
mRNA	messenger RNA
MW	Molecular weigth
NCBI	National Center for Biotechnology Information
NTD	N-terminal domain
OD	Optical density
PAGE	polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
pH	negative logarithm of the hydrogen ion (H^+) concentration
PIC	Preinitiation complex
RNA	Ribonucleic acid
RNAP	RNA polymerase
rpm	revolution per minute

rRNA	ribosomal RNA
RT	room temperature
SDS	sodium dodecylsulphate
SUF13	ORF encoding MBF1
TBP	TATA binding protein
TEMED	N,N,N',N'- Tetramethylethylenediamine
TFB	Transcriptiom factor B, archaeal homologue of TFIIB
TFIIB	Transcriptiom factor II B in Eukaryotes
Tris	Tris-(hydroxymethyl)-aminomethane
tRNA	transference RNA
U	Unit of Enzyme activity
UAS	Upstream activating sequence
UV	Ultraviolet
V	Volt
www	world wide web
g	gravitational acceleration
µ	micro (10^6)

7.2 LIST OF PUBLICATIONS

Marrero, J., Ehrenhofer-Murray, A., Pons, T. and Siebers, B. (2010) Functional analysis of archaeal MBF1 by complementation studies in yeast. Manuscript in preparation.

Marrero, J., Díaz-Valdivia, A. and Coto-Pérez, O. (2010) Molecular mechanisms of heavy metal resistance in bacteria: Applications in bioremediation. **Ciencias Biológicas CENIC**. 41(1): 67-78.

Blombach, F., Makarova, K., Marrero, J., Siebers, B., Koonin, E.V. and van der Oost, J. (2009) Identification of an ortholog of the eukaryotic RNA polymerase III subunit RPC34 in Crenarchaeota and Thaumarchaeota suggests specialization of RNA polymerases for coding and non-coding RNAs in Archaea. **Biol. Direct**, 4:39, doi: 10. 1186/1745-6150-4-39.

Marrero, J., Diaz, A., Valle, A., Cantero, D., Gómez, J.M. and Coto, O. (2009) Nickel and cobalt removal capacities of native metal-resistant bacteria. **Adv. Mat. Research**, 71-73:617-620.

Coto, O., Galizia, F., Hernández, I., Marrero, J. and Donati, E. (2008) Cobalt and nickel recoveries from laterite tailings by organic and inorganic bio-acids. **Hydrometallurgy**, 94:18-22.

Rodríguez-Ramos, T., Espinosa, G., Hernández-López, J., Gollas-Galván, T., Marrero, J., Borrell, Y., Alonso, M.E., Bécquer, U. and Alonso, M. (2008) Effects of *Escherichia coli* lipopolysaccharides and dissolved ammonia on immune response in southern white shrimp *Litopenaeus schmitti*. **Aquaculture**, 274:118–125.

Kouril, T., Zaparty, M., Marrero, J., Klenk, H.P., Brinkmann, H. and Siebers, B. (2008) A novel trehalose synthesizing pathway in the hyperthermophilic Crenarchaeon *Thermoproteus tenax*: The unidirectional TreT pathway. **Arch Microbiol**, 190:355–369.

Marrero, J., Auling, G., Coto, O. and Nies, D. (2007) High-Level Resistance to Cobalt and Nickel in Cuban *Serratia marcescens* strains isolated from serpentine deposits. **Adv. Mat. Research**, 20-21:521-526.

Coto, O., Galizia, F., González, E., Hernández, I., Marrero, J. and Donati, E. (2007) Cobalt and nickel recoveries from laterite tailings by organic and inorganic bioacids. **Adv. Mat. Research**, 20:107–110.

Marrero, J., Auling, G., Coto, O. and Nies, DH. (2007) High-Level Resistance to Cobalt and Nickel but Probably No Transenvelope Efflux: Metal Resistance in the Cuban *Serratia marcescens* strain C-1. **Microb Ecol**, 53(1):123-133.

Coto, O., Peguero, M., Abin, L., Bruguera, N., Marrero J. and Bosecker, K. (2005) Bioleaching of laterite by *Aspergillus niger* strain O-5: an acidophilic, and nickel and cobalt resistant fungus. In: STL Harrison, DE Rawlings, JE Petersen (eds). Bioleaching. **16th Int Biohydromet Symp**, Compress Press, ZA, 357-364.

Marrero J., González, L.J., Castellanos, L., Ayala, M., Fallarero, A. and Coto, O. (2005) Antioxidant defense against cobalt toxicity in *Enterobacter liquefaciens* strain C-1: a highly heavy metal-resistant bacterium. In: STL Harrison, DE Rawlings, J Petersen (eds). Microbial Genetics. **16th Int Biohydromet Symp**, Compress Press, ZA, 781-788.

Marrero J., López, L. J., Sánchez, A., González, W., Ayala, M., Paz-Lago, D., Fallarero, A., Castellanos, L. and Coto, O. (2004) Effect of high concentration of Co (II) on *Enterobacter liquefaciens* strain C-1: A bacterium highly resistant to heavy metals with an unknown genome. **Proteomics**, 4:1265–1279.

Abín, L., Coto, O., Marrero, B. and Marrero, J. (2004) Effect of different factors on citric acid production by *Aspergillus niger* O5. **Ciencias Biológicas CENIC**, 33(3):15-18.

Poster und Vorträge

- Tagung der Vereinigung für Allgemeine und Angewandte Mikrobiologie, VAAM, 2007, Osnabrück (Participant).
- Thermophiles, 2007, Bergen, Norway (Participant).
- DFG Retreat 2007, Role of multiprotein bridging factor 1 (MBF1) as transcriptional co-activator studied in the archaeal model system *Thermoproteus tenax*, Wetter-Volmarstein (Oral presentation).

- Tagung der Vereinigung für Allgemeine und Angewandte Mikrobiologie, VAAM, 2008, Multiple transcription factors (TFBs) in the archaeal transcription of the hyperthermophilic Crenarchaeon *Thermoproteus tenax*, Frankfurt (Poster).
- Promotional Committe Meeting 2008. Function of MBF1 and multiple transcription factors (TFBs) in the archaeal model system *Thermoproteus tenax*, Essen (Oral presentation).
- Molecular Biology of Archaea, 2008. Multiple transcription factors (TFBs) in the archaeal transcription of the hyperthermophilicCrenarchaeon *Thermoproteus tenax*, St Andrews, Scotland (Poster).
- Archaea Meeting, 2008. Transcriptional Regulation in the hyperthermophilic crenarchaeon *Thermoproteus tenax*, Schmitten (Oral presentation).
- DFG Retreat 2008. Function of MBF1 and multiple transcription factors (TFBs) in the archaeal model system *Thermoproteus tenax*, Dormagen Zons (Oral Presentation).
- Tagung der Vereinigung für Allgemeine und Angewandte Mikrobiologie, VAAM, 2009. Multiprotein bridging factor 1 (MBF1) in the archaeal transcription of the hyperthermophilic Crenarchaeon *Thermoproteus tenax* and the mesophilic Euryarchaeon *Methanosarcina mazei*, Bochum (Poster).
- Promotional Committe Meeting 2009. Transcriptional Regulation in the archaeal model system *Thermoproteus tenax*, Essen (Oral presentation).
- DFG Retreat 2009. Transcriptional Regulation in the archaeal model system *Thermoproteus tenax*, Dormagen Zons (Oral Presentation).

7.3 LEBENSLAUF

Persönliche Daten:

Name:	Marrero Coto
Vorname:	Jeannette
Geburtsdatum:	27.09.1977
Geburtsort:	Havana
Staatsangehörigkeit:	Kubanerin
Familienstand:	Verheiratet

Schulausbildung:

1983 - 1989	Primary School, Mariana Grajales, Havana City, Cuba
1989 - 1992	Secondary School, Raul Gómez Garcia, Havana City, Cuba
1992 - 1995	Higher Education in Vocational Institute for Mathematics and Logic Sciences, Havana City, Cuba

Studium:

1995 - 2000	BSc in Biochemistry, University of Havana, Havana, Cuba
2001 - 2003	MSc in Biochemistry and Molecular Biology, University of Havana, Havana, Cuba

Promotion:

2006 - 2009	Promotion an der Duisburg-Essen Universität zum Thema „Transcriptional Regulation in the hyperthermophylic crenarchaeon *Termoproteus tenax*", durchgeführt im Arbeitskreis von Prof. Dr. Bettina Siebers in der Fakultät Chemie, Biofilm Centre, Molekulare Enzymtechnologie und Biochemie der Universität Duisburg-Essen

Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit mit dem Titel

„Transcriptional Regulation in the hyperthermophylic crenarchaeon *Termoproteus tenax*"

selbst verfasst und keine außer den angegebenen Hilfsmitteln und Quellen benutzt habe, und dass die Arbeit in dieser oder ähnlicher Form noch bei keiner anderen Universität eingereicht wurde.

Essen, im April, 2010

UNTERSCHRIFT

Acknowledgements

I would like to express my deep and sincere gratitude to my supervisor, Prof. Dr. Bettina Siebers to accept me as her PhD student, for her support throughout this study, for her constructive comments, for her detailed review and excellent advices during the preparation of this work. Her understanding, encouraging and personal guidance have provided a good basis for the present thesis. It is not sufficient to express my gratitude with only a few words.

I wish to express my warm and sincere thanks to PD. Dr. Winfried Hausner who gave me the opportunity to work in his group in the Archaea Center, University of Regensburg, where I received good training in the footprinting analysis. His kind support, guidance and scientific discussion have been of great value in this study.

I am deeply grateful to Prof. Dr. Ann Ehrenhofer-Murray, to allow me to do all the experiments with yeast in her group in the Department for Genetics (University of Duisburg-Essen), for the scientific discussions and for the interest on my PhD work.

I wish to thank to Dr. Sonja Koning and Fabian Blombach for the experimental contribution previous to the present work.

Many thanks to DFG for the finantial support and all the members of the DFG Research Training Group 1431/1, especially to the organizers, to Prof. Dr. Gerard Rüffel and to the PhD students for the scientific discussion and Ms. Rita Nadorf for her sympathetic help in secretarial work.

I warmly thank Dr. Melanie Zaparty for her support during these years, her valuable advices, her friendly help and for the critical reading of this work.

Very special I want to thank my colleagues from the group in Duisburg for the unforgettable time, for the friendship, timely support and for the nice environment. Thanks: Tina, Mella, Theresa, Patrick, Anna, Sabine, Meike, Julia, Kohei, Dominik and Thomas.

I wish to thank Prof. Hensel to allow me to do the experiments in the isotop laboratory and the members of his group Microbiologie I especially Thomas Knura for his assistance.

Many thanks also to the members of the Department for Genetics (ZMB, Essen) especially to Christhiane Vole and Jan Weber for their generous and timely help during the experimental work.

I am very grateful for Monika Helak for her essential assistance in scanning, sending me the corrected versions and the printing of this work.

There are so many friends who offered me so much kindly help in the difficult moments, especial thanks to Marcia González-Teuber, Familia Acosta-López and Marike Vornkahl.

My special gratitude is due to Family Stephan-Siebers for their loving support. They let me own a happy family in Germany.

I am very grateful for my lovely family in Cuba, especially to my mother who motivates me to work harder and do my best. Last but not least, I am greatly indebted to my devoted husband Yuniel González Gómez. Without his encouragement and understanding it would be impossible for me to finish this work.