

Cloning and functional analysis of the TATA binding protein from *Sulfolobus shibatae*

Sohail A. Qureshi, Peter Baumann, Tracey Rowlands, Bernard Khoo and Stephen P. Jackson*

Wellcome/CRC Institute, Tennis Court Road, Cambridge CB2 1QR, UK and Department of Zoology, Cambridge University, Cambridge, UK

Received January 19, 1995; Revised and Accepted March 28, 1995

GenBank accession no. U23419

ABSTRACT

Archaea (formerly archaeobacteria) comprise a domain of life that is phylogenetically distinct from both Eucarya and Bacteria. Here we report the cloning of a gene from the Archaeon *Sulfolobus shibatae* that encodes a protein with strong homology to the TATA binding protein (TBP) of eukaryotes. *Sulfolobus shibatae* TBP is, however, almost as diverged from other archaeal TBPs that have been cloned as it is from eukaryotic TBPs. DNA binding studies indicate that *S.shibatae* TBP recognizes TATA-like A-box sequences that are present upstream of most archaeal genes. By quantitatively immunodepleting *S.shibatae* TBP from an *in vitro* transcription system, we demonstrate that *Sulfolobus* RNA polymerase is capable of transcribing the 16S/23S rRNA promoter weakly in the absence of TBP. Most significantly, we show that addition of recombinant *S.shibatae* TBP to this immunodepleted system leads to transcriptional stimulation and that this stimulation is dependent on the A-box sequence of the promoter. Taken together, these findings reveal fundamental similarities between the transcription machineries of Archaea and eukaryotes.

INTRODUCTION

Until recently, a long-established dogma in biology was that all life on earth is conveniently divided into two kingdoms: eukaryotes that possess nuclei and prokaryotes that do not. However, 16S rRNA sequence analyses in the 1970s indicated that the prokaryotic world is not a single coherent entity, but, instead, is divided into two distinct groups, the eubacteria and the archaeobacteria (1). Interestingly, these studies indicated that the two groups of prokaryotes are at least as diverged from one another as they are from eukaryotes. Subsequent molecular comparisons of other conserved RNAs and proteins have reinforced this conclusion (2,3). This has led to the proposal to introduce the 'domain' as a new taxonomic level and to divide life on earth into three primary domains: Bacteria, Archaea and Eucarya (3).

Whereas a large amount of research has been directed towards determining the mechanism and regulation of transcription in

Eucarya and Bacteria, relatively little work has been conducted on archaeal transcription systems. Over the past few years, however, several lines of evidence have suggested that the transcription systems of Archaea and Eucarya are fundamentally homologous. Although Archaea possess just one DNA-dependent RNA polymerase, this enzyme consists of ~12 subunits and is therefore similar in its structural complexity to eukaryotic nuclear RNA polymerases (4). In contrast, bacterial RNA polymerases contain three core subunits α , β and β' , and a tightly associated σ factor (5). Immunochemical and sequence analyses have revealed that the two largest RNA polymerase subunits of *Sulfolobus acidocaldarius* are more related to their eukaryotic counterparts than to those of Bacteria (4,6,7). Furthermore, several archaeal RNA polymerase subunits appear to have counterparts in Eucarya, but not in Bacteria (8).

Homologies also exist between archaeal and eukaryotic transcriptional promoter regions. The vast majority of characterized archaeal genes possess an A/T-rich element, the A-box, centred around 27 bp upstream of the transcriptional initiation site. Several studies have indicated that this motif, with a consensus sequence of TTTAWA, has an important role in directing RNA polymerase to the correct transcriptional start site (9,10). This element therefore resembles the TATA-box of eukaryotic RNA polymerase II promoters in both function and sequence. Furthermore, as in eukaryotes, specific and efficient transcriptional initiation from archaeal promoters depends on one or more factors that can be readily separated from the RNA polymerase (11-13). These observations suggest that homologues of eukaryotic general and regulatory transcription factors might exist in Archaea. This speculation has been validated recently by the discovery of a *Pyrococcus woesei* protein that is related to the eukaryotic RNA polymerase II transcription factor TFIIB (14,15) and the identification of TATA-box binding protein (TBP) homologues in *Thermococcus celer* and *P.woesei* (16,17). Studies into the transcriptional machineries of Archaea are therefore likely to have important implications for our understanding of the evolution and functioning of the eukaryotic transcriptional apparatus.

Most evolutionary comparisons indicate that although the Archaea is a monophyletic group, and that it can be divided into two highly diverged kingdoms (3): the Crenarchaeota represent-

* To whom correspondence should be addressed

ing thermophilic sulphur-dependent organisms such as *Sulfolobus* spp. and the Euryarchaeota comprising the methanogens and extreme thermophiles such as *T.celer* and *P.woesei*. Indeed, so divergent are these kingdoms that some analyses have suggested that they do not group together at all and, instead, suggest that the crenarchaeotes might be more related to eukaryotes than are the euryarchaeotes (18,19). In the light of this, it would clearly be of great interest to ascertain whether TBP and/or TFB exist in crenarchaeotes such as *Sulfolobus* and to determine their relatedness to homologues in other systems. The identification of *Sulfolobus* transcription factors would also have practical advantages, because *Sulfolobus* has several features that make it the system of choice for studying transcription in Archaea. First, unlike many Archaea, *Sulfolobus* can be grown easily under aerobic conditions. Secondly, a *Sulfolobus in vitro* transcription system has been established and has been used to characterize in detail the 16S/23S rRNA gene promoter (9,12). Thirdly, the *Sulfolobus* DNA-dependent RNA polymerase has been purified and most of its subunits cloned (4). Fourthly, many *Sulfolobus* genes have already been characterized and several examples of regulated gene expression have been documented (4). In this manuscript we report the cloning of the gene for *S.shibatae* TBP and use the sequence of this factor to investigate the evolutionary relationships of this species to eukaryotes and other Archaea. In addition, we use *in vitro* DNA binding and transcription assays to analyse the function of TBP in the archaeal transcriptional apparatus.

MATERIALS AND METHODS

Polymerase chain reactions (PCR)

Degenerate primers 316 (5'-TAYGARCCIGARCARTTYCC-NGG-3') and 318 (5'-ACRATYTTICISWNSWRAA-3') were used to amplify a segment of the *S.shibatae* TBP gene using low stringency PCR as described previously (17).

Construction and screening of a *S.shibatae* genomic library

The predigested λ -ZAP Express *Bam*HI/CIAP vector cloning kit and packaging extracts were purchased from Stratagene. *S.shibatae* genomic DNA was partially digested with *Sau*3A and fragments ranging in size from 0.5 to 2.5 kb were ligated to the arms and packaged according to the manufacturer's instructions. Screening of the library was performed according to standard procedures (20). Plasmids from the recombinant phage were excised using the ExAssist system (Stratagene).

DNA sequencing and sequence analyses

Plasmid DNAs were purified by centrifugation on CsCl₂ gradients and were sequenced with Sequenase version 2.0 (US Biochemicals) according to the manufacturer's instructions. The sequence of both strands was determined using primer walking. The accession number for the *S.shibatae* TBP sequence is U23419 (GenBank). Database searches for homologous DNA and protein sequences were performed using the FASTA program and the electronic mail-based BLAST server (21,22). Sequence alignments were performed with the UWGCG package using the BESTFIT, PILEUP and LINEUP programs (23) and the ClustalW (1.4) package (24). Phylogenetic trees were computed

using ClustalW and PHYLIP v 3.5 (J. Felsenstein, 1993, Phylogeny Inference Package version 3.5, distributed by the author; Department of Genetics, University of Washington, Seattle, WA) and displayed using TreeTool 2.0.1 (M. Maciukenas, 1994, distributed by the author; Ribosomal RNA Database Project, University of Illinois).

Expression of recombinant *S.shibatae* TBP and generation and use of anti-TBP antisera

Oligonucleotides corresponding to the 5'- and 3'-ends of the *S.shibatae* TBP gene were used in the PCR and the resulting DNA fragment was cloned into the *Bam*HI and *Hind*III sites of the expression vector pQE30 (Qiagen). Bacterially expressed protein was purified to near homogeneity according to manufacturer's instructions on a Ni²⁺-NTA column. For antibody production, purified *S.shibatae* TBP was injected into rabbits five times at 4 week intervals. For Western blot analyses, proteins were transferred from SDS-polyacrylamide gels onto nitrocellulose membranes, probed with antibodies and detected using the ECL system (Amersham).

S.shibatae growth and preparation of extracts

S.shibatae was grown aerobically at 78°C under continuous shaking in a medium recommended by the German Collection of Microorganisms (Sammlungs Katalog 1993). Cells were grown to an OD₅₄₀ of 1.0 prior to harvesting. Cellular extracts were prepared according to Hüdephl *et al.* (12).

Immunodepletion of TBP from *S.shibatae* extracts

Protein-A Sepharose beads (100 μ l) were washed in phosphate-buffered saline, then 100 μ l anti-TBP antiserum was added and the mixture left on ice for 1 h with occasional mixing. The beads were then washed thoroughly with transcription buffer and subsequently incubated on ice for 3 h with 150 μ l *S.shibatae* cell extract (1.5 mg total protein). After centrifugation, the extract was removed, frozen in liquid N₂ and stored at -70°C.

Construction of templates, *in vitro* transcription and DNA binding assays

Based on the *S.shibatae* 16S/23S rRNA gene promoter, two double-stranded oligonucleotides, one containing the A-box and initiator (yR) and the other containing only the yR sequence (see Fig. 3A) were cloned into the *Eco*RV site of pBluescript II SK (Stratagene) to yield the plasmids pAB-4 and pAB-2 respectively. *In vitro* transcription reactions were carried out essentially as described (12) using 0.5 mM rNTPs and 15 μ g *S.shibatae* cell extract. Subsequently, 50 μ l stop buffer (0.1 M Tris-HCl, pH 8.0, 20 mM EDTA, 2% SDS) and 10 μ g *Torula* RNA (Sigma) were added to each reaction, samples extracted with phenol/chloroform and the nucleic acids precipitated from the aqueous phase with ethanol. The probe, binding and gel running conditions for electrophoretic mobility shift assays (EMSAs) were as described previously (17).

Primer extension analysis of *in vitro* synthesized RNA

T7 sequencing primer was end-labelled with polynucleotide kinase (Boehringer Mannheim) in the presence of [γ -³²P]ATP (NEN) and was then separated from unincorporated radiolabel by

passing through a Sephadex G-25 spin column. *In vitro*-synthesized RNA precipitate derived from an *in vitro* transcription assay was resuspended in 10 µl 100 mM Tris-HCl, pH 8.0, 16 mM MgCl₂, 60 mM KCl, 2 mM DTT, 0.2 pmol labelled T7 primer was added and the sample was incubated at 40°C for 25 min. Subsequently, 10 µl extension mix, containing 0.5 mM dNTPs and 25 U avian myeloblastosis virus (AMV) reverse transcriptase (Boehringer Mannheim), was added and the sample incubated for 30 min at 40°C. Nucleic acids were then precipitated using ethanol, washed, dried, resuspended in 50% formamide loading buffer and resolved on a 10% denaturing polyacrylamide gel.

RESULTS

Cloning the gene for a *S.shibatae* TBP homologue

To clone the gene encoding the putative TBP homologue of *S.shibatae*, primers corresponding to regions of TBP that are conserved between eukaryotes and *P.woesei* were synthesized and different combinations were employed in degenerate low stringency PCR reactions using *S.shibatae* genomic DNA. Of the permutations tested, only one primer pair yielded a product of the correct size (data not shown). Subsequent cloning and sequencing of this fragment identified a sequence with the potential to encode a protein with homology to TBP.

Using a region derived from this DNA fragment as a probe to screen a *S.shibatae* genomic library, five positive plaques were identified and subsequently isolated. Sequencing the insert from one of these clones revealed an open reading frame capable of encoding a protein with strong similarity to TBPs characterized previously. As is the case with several other archaeal genes, there is no ATG codon positioned appropriately to serve as the translational start site. In these cases, previous studies have revealed that closely related codons are employed. Assuming that the protein begins with an isoleucine encoded by an ATC codon, *S.shibatae* TBP comprises 198 amino acid residues, has a molecular weight of 22.3 kDa and a pI of 5.9. The DNA sequence of the gene has been deposited in the GenBank database.

In Figure 1A, *S.shibatae* TBP is presented in an alignment with TBP sequences from *P.woesei*, *T.celer* and the C-terminal domains of TBPs from several eukaryotes (not shown are the seven N-terminal amino acid residues ISNSAVS of *S.shibatae* TBP). From this it is clear that the *S.shibatae* protein displays high conservation throughout the core TBP domain, encompassing the two imperfect direct repeat elements. Strikingly, certain residues are identical in every TBP, indicating that they play critical roles in defining the structure of TBP and/or in specifying its functional properties. For example, of the 25 residues of yeast TBP that contact DNA when this protein is bound to the TATA-box motif (25), 23 are identical in *S.shibatae* TBP. Importantly, analysis of the *Sulfolobus* TBP sequence suggests that it contains the structural elements found in eukaryotic TBPs. Taken together, these findings suggest strongly that the *S.shibatae* TBP homologue will bind TATA-related sequences and function in transcription in a manner very similar to its eukaryotic counterparts.

Evolutionary relationships predicted from TBP sequence comparisons

In the light of the controversy surrounding the evolutionary status of *Sulfolobus* and related species with respect to other archaeobacteria

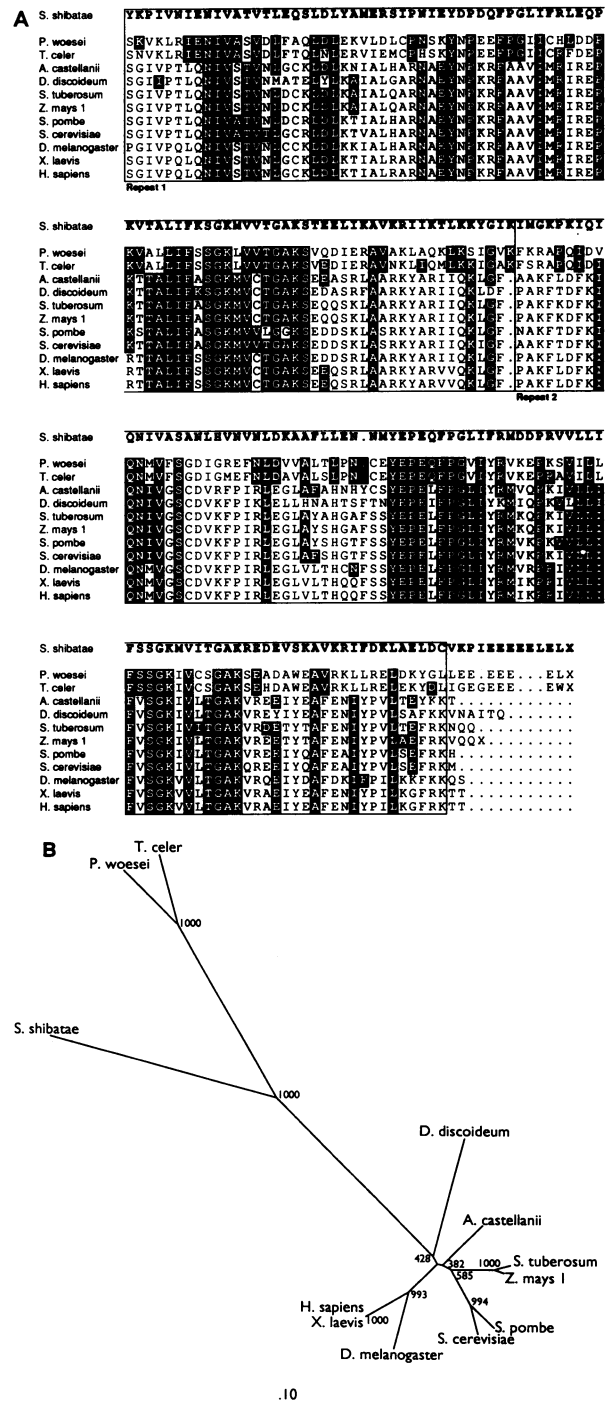


Figure 1. Sequence comparisons of archaeal and eukaryotic TBPs. (A) Alignment of *S.shibatae* TBP with TBPs of eukaryotes and other Archaea. Alscript output (45) of a ClustalW (1.4) alignment of the conserved C-terminal domains of a selection of TBPs. The *S.shibatae* TBP sequence is highlighted in bold type; residues in the other TBPs which are identical to the corresponding *S.shibatae* residue are set in reverse shading. Accession numbers for the sequences used are: *S.shibatae* U23419, *P.woesei* U10285; *T.celer* U04932; *Acanthamoeba castellanii* P26354; *Dictyostellium discoideum* P26355; *Solanum tuberosum* P26357; *Zea mays* TBP1 L13301; *Schizosaccharomyces pombe* P17871; *Saccharomyces cerevisiae* P13393; *Drosophila melanogaster* P20227; *Xenopus laevis* P27633; *Homo sapiens* P20226. (B) Phylogenetic tree based on TBP sequence comparisons. Radial plot of a phylogenetic tree inferred from the sequence alignments shown in (A). The numbers next to each node refer to the number of bootstrap trials (out of 1000) which support the branching pattern at each node. Scale bar, 10 mutations/100 sequence positions.

Table 1. Comparison of *S.shibatae* TBP with TBPs of eukaryotes and other Archaea

	<i>S. shibatae</i>	<i>P. woesei</i>	<i>T. celer</i>	<i>A. castellanii</i>	<i>D. discoideum</i>	<i>S. tuberosum</i>	<i>Z. mays 1</i>	<i>S. pombe</i>	<i>S. cerevisiae</i>	<i>D. melanogaster</i>	<i>X. laevis</i>	<i>H. sapiens</i>
<i>S. shibatae</i>		0.963	0.918	1.120	1.209	1.108	1.105	1.139	1.070	1.238	1.202	1.202
<i>P. woesei</i>	0.453		0.173	1.312	1.490	1.328	1.327	1.401	1.289	1.296	1.339	1.339
<i>T. celer</i>	0.469	0.849		1.264	1.399	1.308	1.291	1.348	1.243	1.231	1.276	1.276
<i>A. castellanii</i>	0.425	0.380	0.385		0.231	0.133	0.133	0.179	0.157	0.196	0.169	0.169
<i>D. discoideum</i>	0.397	0.341	0.352	0.810		0.258	0.243	0.281	0.293	0.312	0.291	0.291
<i>S. tuberosum</i>	0.430	0.374	0.369	0.883	0.793		0.012	0.153	0.172	0.231	0.187	0.187
<i>Z. mays 1</i>	0.430	0.374	0.374	0.883	0.804	0.972		0.135	0.147	0.225	0.181	0.181
<i>S. pombe</i>	0.419	0.369	0.374	0.844	0.771	0.860	0.877		0.075	0.239	0.232	0.232
<i>S. cerevisiae</i>	0.436	0.391	0.397	0.860	0.760	0.849	0.866	0.933		0.232	0.221	0.221
<i>D. melanogaster</i>	0.408	0.385	0.385	0.832	0.749	0.816	0.821	0.799	0.799		0.133	0.133
<i>X. laevis</i>	0.402	0.374	0.380	0.849	0.760	0.838	0.844	0.799	0.804	0.888		0.000
<i>H. sapiens</i>	0.402	0.374	0.380	0.849	0.760	0.838	0.844	0.799	0.804	0.888	1.000	

The figures comprising the lower left triangle show the fraction of identical amino acids between each pair of TBPs from the alignments depicted in Figure 1. The figures comprising the upper right triangle show the average number of amino acid replacements per position between pairs, calculated using the Dayhoff PAM option in the PROTDIST program (J. Felsenstein; see Materials and Methods).

and eukaryotes, we analysed the relationships between the TBP molecules of these three groups. In several ways the *S.shibatae* protein is more related to the TBPs of *T.celer* and *P.woesei* than to those of eukaryotes. First, all three archaeal TBPs comprise essentially just the core domain and possess very short N-terminal extensions. This contrasts with the appreciable N-terminal regions that are found commonly in eukaryotic TBPs. Secondly, all the archaeal proteins are acidic, having pI values ranging from 4.7 to 5.9, whereas the pI for the core domains of eukaryotic TBP molecules is typically ~10.3. It is tempting to speculate that these features of the archaeal proteins might be adaptations to the high temperature growth conditions of these organisms.

In spite of the similarities between the archaeal TBPs noted above, sequence identity comparisons indicate that the *S.shibatae* protein is only slightly more related to the TBPs of *T.celer* and *P.woesei* than it is to the core domain of eukaryotic TBPs (Table 1). For example, its identities to *P.woesei* and human are 45.3% and 40.2% respectively. Consistent with *T.celer* and *P.woesei* being relatively closely related, the TBPs from these organisms are 84.9% identical. Interestingly, eukaryotic TBPs display a slightly higher degree of sequence identity to *S.shibatae* TBP than to the TBPs of *P.woesei* and *T.celer*. Although this may reflect the fact that *P.woesei* and *T.celer* grow at higher temperatures than *S.shibatae*, these data are consistent with the suggestion that organisms in the crenarchaeotic branch of the Archaea share relatively more features with eukaryotes than those in the euryarchaeotic branch.

To assess more accurately the evolutionary relationships between the various TBPs, the aligned sequences depicted in Figure 1A were used to construct a phylogenetic tree (Fig. 1B). As expected, the eukaryotic TBPs are closely related to each other. Most importantly, this analysis indicates that the archaeal TBP sequences group together. An attempt to root the tree was made based on the assumption that the two direct repeats arose prior to the divergence of Archaea and Eucarya. This analysis reveals that the root does not lie within the archaeal grouping (data not shown), therefore supporting a monophyletic model for the Archaea and suggesting that TBP was already in existence in the last common ancestor of all Eucarya and Archaea. However, although bootstrap analyses indicate that the archaeal grouping is statistically secure, the relationship between TBPs of *S.shibatae*

and the other Archaea is only marginally greater than the relationship between *S.shibatae* and eukaryotic TBP proteins.

Since the two related direct repeat elements of TBP are believed to have arisen through a duplication event, it is enlightening to compare the degree of divergence between the two repeats within various TBP molecules. It is of interest to note that, whereas the fraction of identical amino acid residues between the two repeats of the core domain of human TBP is only 0.27, this value is 0.39 for *P.woesei* TBP and 0.48 for *S.shibatae* TBP (data not shown). Consistent with the strong relationship between the two halves of archaeal TBPs, many archaeal A-box elements correspond to the perfect palindromic sequence TTTAAA (4), suggesting that the two domains of archaeal TBPs may be functionally equivalent. One interpretation for the higher relatedness between the two repeats of archaeal TBPs than between the repeats of eukaryotic TBPs is that the duplication(s) that generated the archaeal TBPs occurred more recently than that which generated eukaryotic TBPs. This proposal is not, however, supported by phylogenetic analyses conducted with the two halves of the *S.shibatae* sequence (data not shown), which support the contention that a single duplication event occurred in the common Archaeal/Eucaryal ancestor. Instead, these results are consistent with studies on early gene duplications, which have suggested that the evolutionary rate is lower in most Archaea than in Eucarya (for example 26). Archaeal TBPs, and *S.shibatae* TBP in particular, may therefore resemble closely the TBP molecule present in the last common ancestor of all Archaea and Eucarya.

S.shibatae TBP binds to archaeal A-box elements

The striking homologies between the TBP molecules of *S.shibatae* and eukaryotes suggests that the *S.shibatae* protein will interact with the TATA-related A-box elements found upstream of most archaeal genes. To test this prediction, we purified *Escherichia coli*-expressed TBP to virtual homogeneity (Fig. 2A) and utilized it in an electrophoretic mobility shift assay (EMSA) using an oligonucleotide containing the A-box sequence (TTTA-TAT) of the *S.shibatae* 16S/23S rRNA gene promoter. As can be seen from Figure 2B, when recombinant TBP is incubated with this oligonucleotide, a specific protein-DNA complex is formed (lanes 2-4). A control fraction derived from *E.coli* containing the

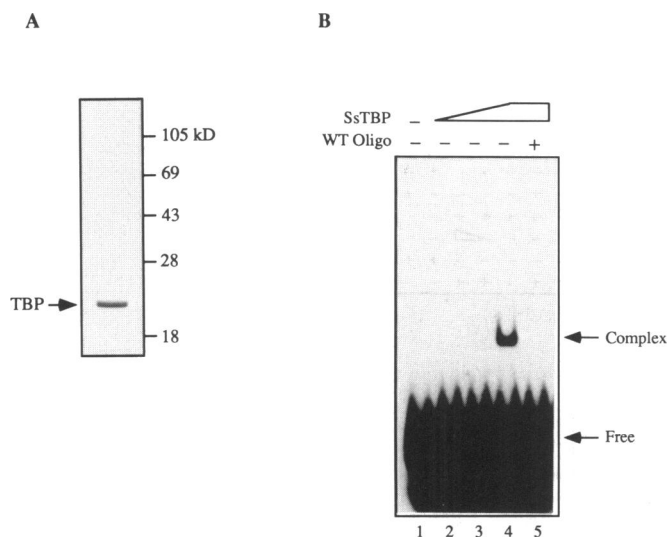


Figure 2. *S.shibatae* TBP binds to archaeal A-box elements. (A) Purity of the bacterially expressed recombinant factors. Histidine-tagged *S.shibatae* TBP was electrophoresed on a 12% SDS-polyacrylamide gel and detected by Coomassie staining. The location of this protein and protein size markers are indicated. (B) EMSA assays were conducted using a radiolabelled oligonucleotide (0.1 ng) containing an A-box sequence as described in Materials and Methods, using 0, 2, 20 and 200 ng *S.shibatae* (Ss) TBP (lanes 1–4). Reaction 5 received 200 ng SsTBP. Assays were conducted in the absence (lanes 1–4) or presence (lane 5) of wild-type (WT) A-box-containing oligonucleotide competitor (10 ng). The positions of free DNA and TBP–DNA complex are indicated.

parental expression vector does not yield this complex (data not shown). More importantly, the specificity of the interaction was established by the fact that the *S.shibatae* TBP–DNA complex can be competed effectively by addition of excess unlabelled A-box-containing oligonucleotide (lane 5).

The role of TBP in *S.shibatae* 16S/23S rDNA transcription

Since TBP is an essential component of all three eukaryotic nuclear transcription systems (for reviews see 27,28), we speculated that *S.shibatae* TBP would play an important role in *S.shibatae* transcription. To test this prediction, it was necessary to establish a *Sulfolobus in vitro* transcription system. To this end, we constructed two plasmid templates: one, pAB-4, contains the A-box sequence and the initiator (yR) element of the well-characterized strong promoter of the *S.shibatae* 16S/23S rRNA gene; the other plasmid, pAB-2, contains only the yR sequence (Fig. 3A). These plasmids were then employed in *in vitro* transcription studies using *S.shibatae* cell extracts. In order to detect *in vitro*-synthesized transcripts and to map the site of transcriptional initiation, we used a 17 base oligonucleotide (T7 sequencing primer) in a primer extension assay with the *in vitro*-transcribed RNA. Importantly, and consistent with previous observations (12), these experiments revealed that high levels of transcription are attained in the *Sulfolobus in vitro* system and that the transcriptional start site employed matches the initiation site for the 16S/23S rRNA precursor *in vivo* (the initiating guanine residue is indicated by an arrow in Fig. 3A).

In order to determine whether TBP plays a role in transcription in the *S.shibatae* system and to ascertain whether such an effect is mediated by it interacting with the A-box sequence, we used

anti-TBP antisera to deplete TBP from the *S.shibatae* extract. Whereas pre-immune sera are unable to remove TBP from *S.shibatae* cell extracts (Fig. 3B, lane 2), essentially all of this protein is immunodepleted when an anti-*S.shibatae* TBP antiserum is used (lane 3). We then compared this depleted extract with undepleted and mock-depleted controls for its ability to direct transcription *in vitro*. For these studies we employed pAB-4 that contains the A-box and initiator elements of the 16S/23S rRNA promoter and its derivative pAB-2 that lacks the A-box sequence. As can be seen from Figure 3C, the untreated *S.shibatae* extract transcribes pAB-4 much more strongly than pAB-2 (lanes 1 and 6). This is consistent with previous reports that archaeal RNA polymerases can weakly recognize promoters that lack A-box elements (12). Mock-depletion of the extract does not lead to reduced transcription of either pAB-4 or pAB-2 (lanes 2 and 7), indicating that the immunoprecipitation incubation conditions do not, in themselves, inactivate any essential transcription factor. Significantly, immunodepletion of TBP was found to have no effect on transcription of pAB-2, which lacks the A-box element (Fig. 3C, lane 8), but led to a dramatic reduction in transcription of pAB-4, which contains the A-box (lane 3). Consistent with this decreased transcription of pAB-4 being directly attributable to the removal of TBP, high levels of transcription are restored upon addition of recombinant *S.shibatae* TBP (lanes 4 and 5). In contrast, TBP addition does not affect transcription of pAB-2 (lanes 9 and 10). Taken together, these results indicate that the *S.shibatae* transcriptional apparatus can recognize promoters that lack A-box elements weakly and that this low level of transcription is independent of TBP. Furthermore, they establish that *S.shibatae* TBP is indeed a *bona fide* archaeal transcription factor and that it functions through A-box sequences that are found upstream of most archaeal genes.

DISCUSSION

In this manuscript we have investigated the transcriptional apparatus of the archaeon *S.shibatae* by cloning the gene for its TBP homologue. Comparison of the *S.shibatae* TBP sequence with other TBP sequences strongly supports, but does not prove, a monophyletic model for the Archaea. However, there is a striking degree of divergence between *S.shibatae* (and, by inference, its crenarchaeotic relatives) and Euryarchaeota such as *Pwoesei* and *T.celer*. This suggests that the archaeal lineage split into the two branches relatively shortly after the common archaeal ancestor diverged from the eukaryotic line of descent.

To study the transcriptional properties of *S.shibatae* TBP and as a step towards defining in detail the mechanism of transcription in Archaea we have used highly specific anti-TBP antisera to immunodeplete TBP from *S.shibatae* extracts. When these extracts are used to transcribe templates with or without the TATA-like A-box sequence, equally low levels of transcription are obtained. The fact that addition of TBP stimulates transcription from only the A-box-containing promoter illustrates that TBP mediates its stimulatory effect by interacting with the A-box. Taken together, therefore, the available data indicate that there are striking similarities between the mechanisms by which TBP functions in Archaea and Eucarya. In each case, TBP proteins recognize TATA-box-like sequences and then direct transcription to initiate 25–30 bp downstream. This contrasts with transcription in Bacteria, where no clear TBP homologues have been detected

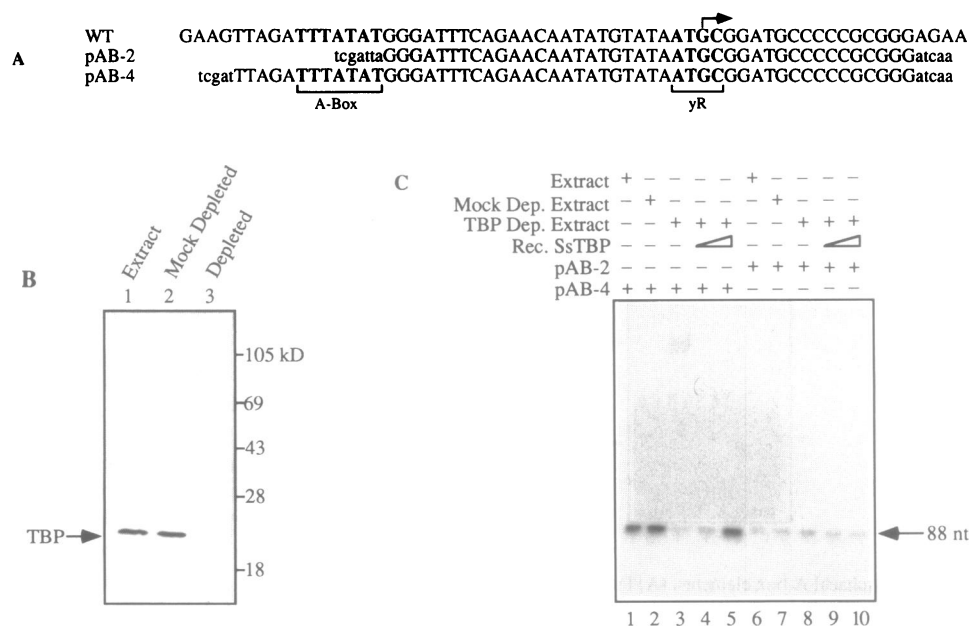


Figure 3. *S. shibatae* TBP stimulates transcription by interacting with the TATA-like A-box sequence. (A) DNA sequence of the *S. shibatae* 16S/23S promoter region (WT) and of the 16S/23S rRNA promoter fragments cloned into pBluescript to yield plasmids pAB-2 and pAB-4 (upper case letters). Plasmid DNA sequence is shown in lower case letters. Transcription initiation site in the initiator (yR) sequence is shown by an arrow. (B) Western blot analysis of the immunodepleted extract. Total extract (lane 1), mock-depleted extract (lane 2) and TBP-depleted extract (lane 3). The position of the TBP band is shown by an arrow. (C) *In vitro* transcription analysis of pAB-2 and pAB-4 using undepleted and TBP-depleted extracts. Transcription reactions were carried out on pAB-4 (lanes 1–5) and pAB-2 (lanes 6–10) using undepleted extract (lanes 1 and 6), mock-depleted extract (lanes 2 and 7) and TBP-depleted extract (lanes 3–5 and 8–10). Reactions 4 and 9, and 5 and 10, were supplemented with 2 and 20 ng recombinant *S. shibatae* TBP (SsTBP) respectively. The 88 nt primer extension product is shown by an arrow.

and where transcriptional specificity is imparted by σ factors that are normally tightly associated with the RNA polymerase (5,29).

In the light of the similarity between TBP function in Archaea and Eucarya, it is tempting to speculate that other components of the transcriptional machinery might be conserved in structure and in function between these two groups of organisms. Indeed, the cloning of genes for several archaeal RNA polymerase subunits has established that they bear strong resemblance to subunits of eukaryotic RNA polymerases I–III (4). Consistent with our finding that *S. shibatae* RNA polymerase is able to transcribe the 16S/23S rRNA promoter weakly even in the absence of TBP, previous studies have revealed that purified archaeal RNA polymerases have an intrinsic ability to recognize certain promoters (30,31). This suggests that promoter recognition by the RNA polymerase may be mediated by direct interaction with the weakly conserved B-box sequence that spans the initiation sites of several archaeal genes (12). Because of the similarities between archaeal and eukaryotic RNA polymerases, it is tempting to speculate that eukaryotic RNA polymerases might interact with initiator elements that have been defined in several class II promoters (32). Alternatively, it is possible that eukaryotic factors that have been implicated in binding initiator elements are evolutionarily related to protein(s) associated with the archaeal RNA polymerase.

In eukaryotes, TBP is found complexed with TBP-associated factors (TAFs; 33,34). For example, TBP associates with different sets of TAFs to form the complexes SL1, TFIID and TFIIB that are dedicated to the RNA polymerase I, II and III systems respectively (for reviews see 27,28). Since eukaryotic

TAFs appear to play important roles in receiving regulatory stimuli (for example see 35,36) and in mediating recognition of certain promoters (for example see 37,38), it is tempting to speculate that multiprotein TBP–TAF complexes might also exist in Archaea. Although this may be the case, unlike the situation in eukaryotes, where TBP–TAF complexes are generally >200 kDa in size, we find that *S. shibatae* TBP migrates on glycerol gradients in a manner consistent with a molecular weight of ~50 kDa (data not shown). This might correspond to a TBP dimer or might represent *S. shibatae* TBP associated with a small number of TAF polypeptides.

In the eukaryotic RNA polymerase II system, several factors in addition to TBP and RNA polymerase are required for efficient transcription. One factor that is particularly crucial is TFIIB, which stabilizes the binding of TBP to TATA-box elements and serves as a molecular bridge between TBP and RNA polymerase II (for example see 39,40). Similarly, the TFIIB-related factor BRF is required to recruit RNA polymerase III to class III genes (41–43). A TFIIB/BRF-related protein (TFB) has been cloned from *P. woesei* (14,15) and shown to potentiate the association of *P. woesei* TBP with archaeal A-box elements (17). Furthermore, we have recently isolated the gene for *S. shibatae* TFB and have shown that this protein increases the rate of assembly and/or stability of complexes between *S. shibatae* TBP and DNA (44). With the availability of a highly efficient *S. shibatae in vitro* transcription system, it should be possible to ascertain the precise roles played by TBP, TFB and RNA polymerase and to identify other components of the archaeal transcriptional apparatus.

ACKNOWLEDGEMENTS

We thank B. Cherolambus and H. Connaris for a starter culture of *S.shibatae* and advice on growing conditions and Graeme Smith for help with the sequence alignments. Finally, we thank members of the SPJ laboratory for advice and valuable discussions. PB is currently at the ICRF Clare Hall Laboratories, South Mimms, Herts, UK. This work was funded by grant SP2143/0101 from the Cancer Research Campaign and by grant G9420496MB from the Medical Research Council (UK).

REFERENCES

- 1 Woese,C.R. and Fox,G.E. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5088–5090.
- 2 Fox,G.E. *et al.* (1980) *Science*, **209**, 457–463.
- 3 Woese,C.R., Kandler,O. and Wheelis,M.L. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 4576–4579.
- 4 Zillig,W., Palm,P., Klenk,H.-P., Langer,D., Hüdepohl,U., Hain,J., Lanzendorfer,M. and Holz,I., (1993) In Kates,M., Kushner,D.J. and Matheson,A.T. (eds), *The Biochemistry of Archaea (Archaeobacteria)*. Elsevier, New Comprehensive Biochemistry, Vol. 26, pp. 367–391.
- 5 Eick,D., Wedel,A. and Heumann,H. (1994) *Trends Genet.*, **10**, 292–296.
- 6 Berghofer,B., Krockel,L., Kortner,C., Truss,M., Schallenberg,J. and Klein,A. (1988) *Nucleic Acids Res.*, **16**, 8113–8128.
- 7 Pühler,G., Lottspeich,F. and Zillig,W. (1989) *Nucleic Acids Res.*, **17**, 4517–4534.
- 8 Klenk,H.-P., Palm,P., Lottspeich,F. and Zillig,W. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 407–410.
- 9 Reiter,W.-D., Hüdepohl,W. and Zillig,W. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 9509–9513.
- 10 Hausner,W., Frey,G. and Thomm,M. (1991) *J. Mol. Biol.*, **222**, 495–508.
- 11 Frey,G., Thomm,M., Brudigam,B., Gohl,H.P. and Hausner,W. (1990) *Nucleic Acids Res.*, **18**, 1361–1367.
- 12 Hüdepohl,U., Reiter,W.-D. and Zillig,W. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 5851–5855.
- 13 Hausner,W. and Thomm,M. (1993) *J. Biol. Chem.*, **268**, 24047–24052.
- 14 Ouzounis,C. and Sander,C. (1992) *Cell*, **71**, 189–190.
- 15 Creti,R., Londei,P. and Cammarano,P. (1993) *Nucleic Acids Res.*, **21**, 2942.
- 16 Marsh,T.L., Reich,C.I., Whitelock,R.B. and Olsen,G.J. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 4180–4184.
- 17 Rowlands,T., Baumann,P. and Jackson,S.P. (1994) *Science*, **264**, 1326–1329.
- 18 Lake,J.A. (1988) *Nature*, **331**, 184–186.
- 19 Rivera,M.C. and Lake,J.A. (1992) *Science*, **257**, 74–76.
- 20 Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (1991) In Janssen,K. (ed.), *Current Protocols in Molecular Biology*. J.Wiley & Sons.
- 21 Pearson,W.R. and Lipman,D.J. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 2444–2448.
- 22 Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipmann,D. (1990) *J. Mol. Biol.*, **215**, 403–410.
- 23 Deveraux,J., Haerberli,P. and Smithies,O. (1984) *Nucleic Acids Res.*, **12**, 387–395.
- 24 Thompson,J.D., Higgins,D.G. and Gibson,T.J. (1994) *Nucleic Acids Res.*, **22**, 4673–4680.
- 25 Kim,Y., Geiger,J.H., Hahn,S. and Sigler,P.B. (1993) *Nature*, **365**, 512–520.
- 26 Iwabe,N., Kuma,K.-I., Hasegawa,M., Osawa,S. and Miyata,T. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 9355–9359.
- 27 White,R.J. and Jackson,S.P. (1992) *Trends Genet.*, **8**, 284–288.
- 28 Hernandez,N. (1993) *Genes Dev.*, **7**, 1291–1308.
- 29 Busby,S. and Ebright,R.H. (1994) *Cell*, **79**, 743–746.
- 30 Thomm,M. and Wich,G. (1988) *Nucleic Acids Res.*, **16**, 151–157.
- 31 Brown,J.W., Thomm,M., Beckler,G.S., Frey,G., Setter,K.O. and Reeve,J.N. (1988) *Nucleic Acids Res.*, **16**, 135–150.
- 32 Smale,S.T. and Baltimore,D. (1989) *Cell*, **57**, 103–113.
- 33 Goodrich,J.A. and Tjian,R. (1994) *Curr. Opin. Cell Biol.*, **6**, 403–409.
- 34 Poon,D. and Weil,P.A. (1993) *J. Biol. Chem.*, **268**, 15325–15328.
- 35 Chen,J.-L., Attardi,L.D., Verrijzer,C.P., Yokomori,K. and Tjian,R. (1994) *Cell*, **79**, 93–105.
- 36 White,R.J., Gottlieb,T.M., Downes,C.S. and Jackson,S.P. (1995) *Mol. Cell Biol.*, **15**, 1983–1992.
- 37 Kaufmann,J. and Smale,S.T. (1994) *Genes Dev.*, **8**, 821–829.
- 38 Purnell,B.A., Emanuel,P.A. and Gilmour,D.S. (1994) *Genes Dev.*, **8**, 830–842.
- 39 Ha,I., Roberts,S., Maldonado,E., Sun,X., Kim,L.-U., Green,M. and Reinberg,D. (1993) *Genes Dev.*, **7**, 1021–1032.
- 40 Yamashita,S., Hisatake,K., Kokubo,T., Doi,K., Roeder,R.G., Horikoshi,M. and Nakatani,Y. (1993) *Science*, **261**, 463–466.
- 41 Kassavetis,G.A., Joazeiro,C.A.P., Pisano,M., Geiduschek,E.P., Colbert,T., Hahn,S. and Blanco,J.A. (1992) *Cell*, **71**, 1055–1064.
- 42 Werner,M., Chaussivert,N., Willis,I.M. and Sentenac,A. (1993) *J. Biol. Chem.*, **268**, 20721–20724.
- 43 Khoo,B., Brophy,B. and Jackson,S.P. (1994) *Genes Dev.*, **8**, 2879–2890.
- 44 Qureshi,S.A., Khoo,B.C.E., Baumann,P. and Jackson,S.P. (1995) *Proc. Natl. Acad. Sci. USA*, in press.
- 45 Barton,G.J. (1993) *Protein Engng*, **6**, 37–40.