The EMBO Journal Vol.16 No.10 pp.2927–2936, 1997

Factor requirements for transcription in the Archaeon *Sulfolobus shibatae*

Sohail A.Qureshi, Stephen D.Bell and Stephen P.Jackson[1]

Wellcome/CRC Institute, Tennis Court Road, Cambridge CB2 1QR and Department of Zoology, University of Cambridge, Cambridge, CB2 3EJ, UK

[1]Corresponding author at Wellcome/CRC Institute

Archaea (archaebacteria) constitute a domain of life that is distinct from Bacteria (eubacteria) and Eucarya (eukaryotes). Although archaeal cells share many morphological features with eubacteria, their transcriptional apparatus is more akin to eukaryotic RNA polymerases I, II and III than it is to eubacterial transcription systems. Thus, in addition to possessing a 10 subunit RNA polymerase and a homologue of the TATA-binding protein (TBP), Archaea possess a polypeptide termed TFB that is homologous to eukaryotic TFIIB. Here, we investigate the factor requirements for transcription of several promoters of the archaeon *Sulfolobus shibatae* and its associated virus SSV. Through *in vitro* transcription and immunodepletion, we demonstrate that *S.shibatae* TBP, TFB and RNA polymerase are not complexed tightly with one another and that each is required for efficient transcription of all promoters tested. Furthermore, full transcription is restored by supplementing respective depleted extracts with recombinant TBP or TFB, indicating that TBP-associated factors or TFB-associated factors are not required. Indeed, gel-filtration suggests that *Sulfolobus* TBP and TFB are not associated stably with other proteins. Finally, all promoters analysed are transcribed accurately and efficiently in an *in vitro* system comprising recombinant TBP and TFB, together with essentially homogeneous preparation of RNA polymerase. Transcription in Archaea is therefore fundamentally homologous to that in eukaryotes, although factor requirements appear to be much less complex.

Keywords: Archaea/RNA polymerase/TBP/TFB/transcription

Introduction

Life has been traditionally divided into two domains—eukaryotes and bacteria. However, analyses of 16S/23S rRNA sequences by Woese and colleagues in the late 1970s led to the proposal that the prokaryotic world is actually divided into two distinct domains, which are at least as diverged evolutionarily from one another as they are from eukaryotes (Woese and Fox, 1977). Subsequent studies on other conserved cellular RNAs and proteins substantiated this view (e.g. Fox *et al.*, 1980; Woese *et al.*, 1990), and it is now generally accepted that life on earth is divided into three domains termed Eucarya (eukaryotes), Bacteria (eubacteria) and Archaea (archaebacteria). Although many characterized archaeal species are capable of thriving at high temperatures and at extremes of salinity and/or pH, this does not appear to be a general feature of these organisms. For example, it has been shown that Archaea account for up to 30% of picoplankton in certain Antarctic waters (DeLong *et al.*, 1994). In addition, recent molecular sampling studies have led to the identification of a hitherto unsuspected diversity of archaeal species (Barns *et al.*, 1996). In light of these points, it is clear that Archaea play very important roles in the biosphere, and it is therefore of great importance to achieve a better understanding of their biology and molecular organization.

Unlike the situation for Eucarya and Bacteria, relatively little is known about the mechanism and regulation of gene expression in Archaea. However, work over the past decade has indicated that the archaeal transcription apparatus is more similar to the RNA polymerase (pol) I, II and III systems of eukaryotic cell nuclei than it is to the eubacterial transcription apparatus (for reviews see Qureshi and Jackson, 1997; Baumann *et al.*, 1995). For example, although Archaea have only one RNA polymerase, it consists of ~10 subunits and is therefore similar in complexity to the eukaryotic nuclear enzymes (Zillig *et al.*, 1993). In contrast, the eubacterial RNA polymerase is simpler in its molecular composition, comprising a σ-factor and three core subunits—α, β and β′. Significantly, immunochemical and sequence analyses have revealed that the two largest archaeal RNA polymerase subunits are more related to their eukaryotic counterparts than to those of Bacteria (Berghofer *et al.*, 1988; Puhler *et al.*, 1989; Zillig *et al.*, 1993). Furthermore, several other archaeal RNA polymerase subunits appear to have counterparts in Eucarya but not in Bacteria (Klenk *et al.*, 1992; Langer *et al.*, 1995). In addition, archaeal promoters resemble those for eukaryotic pol II in that they contain a highly conserved TATA-like A-box sequence (consensus $TTTA^{A}/_{T}A$), located 25–30 bp upstream from the transcriptional initiation site, and also possess a weakly conserved initiator element. Several studies have shown that the A-box element plays an important role in directing the archaeal RNA polymerase to the correct transcription start site (Reiter *et al.*, 1990; Hausner *et al.*, 1991), and accurate transcriptional initiation has been shown to require factors that can be separated readily from the archaeal RNA polymerase (Frey *et al.*, 1990; Hüdepohl *et al.*, 1990; Hausner and Thomm, 1993).

In Bacteria, the only components required for basal transcription are the RNA polymerase core enzyme and a specificity (σ) factor that interacts with the core enzyme and directs it to the promoter DNA (reviewed in Eick *et al.*, 1994). In contrast, eukaryotic RNA polymerases

require multiple, chromatographically separable, accessory factors to direct them to their respective promoters. For example, six distinct general transcription factors—TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH—function in specifying accurate basal transcription by pol II (reviewed in Orphanides *et al*., 1996; Roeder, 1996). In many cases, the first step in pol II transcription complex assembly involves interactions between the TATA-box and TFIID, which is a large multi-subunit complex consisting of the TATA-binding protein (TBP) and at least eight tightly associated factors (TAFs) ranging in size from 18 to 250 kDa (e.g. Tanese *et al*., 1991; Verrijzer and Tjian, 1996). This is followed by the binding of TFIIA and the 35 kDa basal transcription factor TFIIB. Subsequently, TFIIF facilitates the recruitment of pol II to the TFIID–TFIIA–TFIIB preinitiation complex, and this is followed by the sequential binding of TFIIE and TFIIH prior to transcriptional initiation and promoter clearance. In addition to TFIID, eukaryotic cells contain three other TBP–TAF complexes, each of which has its own unique set of TAFs that are distinct from those of TFIID (reviewed in White and Jackson, 1992; Hernandez, 1993). The additional TBP–TAF complexes are the pol I specificity factor SL1, the pol III specificity factor TFIIIB and the U6 snRNA gene specific pol III factor $SNAP_C$. Interestingly, one TFIIIB component is the protein BRF, which is related in sequence to TFIIB. Like TFIIB, BRF plays an important role in recruiting RNA polymerase to the promoter (e.g. Kassavetis *et al.*, 1990; Werner *et al*., 1993; Khoo *et al*., 1994). Although somewhat less complex than the situation for pol II, pols I and III also require basal factors in addition to their respective TBP–TAF complexes. Thus, pol I utilizes the protein UBF (Hempel *et al*., 1996), transcription of most pol III promoters requires the multiprotein complex TFIIIC, and transcription of the 5S rRNA gene by pol III requires TFIIIA (White, 1994).

The presence of TATA-box like sequences in most archaeal promoters, together with the similarities between the archaeal and eukaryotic RNA polymerases, suggests that the mechanisms of transcription may be similar in these two groups of organisms. Consistent with this, a TBP homologue has been identified in several archaeal species (Marsh *et al*., 1994; Rowlands *et al*., 1994; Qureshi *et al*., 1995a; Bult *et al.*, 1996), and it has been shown that this factor binds specifically to the A-box motif (Rowlands *et al*., 1994; Qureshi *et al*., 1995a). In addition, a homologue of TFIIB/BRF has been identified in several species of Archaea (Ouzounis and Sander, 1992; Creti *et al*., 1993; Qureshi *et al*., 1995b; Bult *et al.*, 1996), and the *Sulfolobus shibatae* and *Pyrococcus woesei* proteins have been shown to stimulate the binding of their cognate TBP to the A-box element. Since Archaea contain only one RNA polymerase, we refer to the TFIIB homologue as TFB (Qureshi *et al*., 1995b). It is not currently known whether Archaea possess homologues of other eukaryotic basal transcription factors, or whether different archaeal promoters have distinct basal factor requirements. To address these issues, we have used a combination of immunochemical and biochemical approaches to analyse the involvement of *S.shibatae* TBP and TFB in the transcription of several promoters derived from *S.shibatae* and its associated virus SSV (Schleper *et al*., 1992). In addition, we have tested whether or not TBP and TFB are complexed relatively stably with one another, with the RNA polymerase or with other proteins. Finally, through using a highly defined *Sulfolobus in vitro* transcription system, we have addressed the question of whether there exist other archaeal basal factors. Together, these studies provide important insights into the mechanism of transcriptional initiation in Archaea and its relationship to transcription in Eucarya.

Results

Accurate and efficient transcription by crude S.shibatae cell extracts

To initiate our analyses of archaeal transcription, we first cloned the *S.shibatae* TFB and 16S/23S rRNA promoters, together with the SSV T5 and T6 promoters (Reiter *et al*., 1988), into plasmid vectors. The resulting constructs were then tested for their ability to be transcribed using a crude extract of *S.shibatae* cells. In each case, the *in vitro* transcription products were detected by primer extension analysis using an oligonucleotide based on downstream plasmid sequences, and the precise site(s) of transcriptional initiation was determined by electrophoresing the primer extension product alongside its respective DNA sequencing reactions (Figure 1A). This showed that each of the four promoters initiates transcription *in vitro* with high efficiency and specificity. Moreover, the sites for transcriptional initiation map precisely to those employed *in vivo* (Reiter *et al*., 1987, 1988; Qureshi *et al*., 1995b) and each initiation site lies 25–30 bp downstream from a TATA-like A-box element (Figure 1B). Although no convincing consensus initiator sequence has been established for archaeal promoters, one feature of the initiation sites employed by the four promoters analysed here is that they are all located one nucleotide upstream from a guanine residue. Whereas the T5, T6 and 16S/23S promoters each produce a single major transcription product, two predominant sites for initiation are observed for the TFB promoter *in vivo* (Qureshi *et al*., 1995b) and *in vitro* (Figure 1), which is consistent with the fact that this promoter contains two A-box elements separated by 24 bp. Significantly, the upstream TFB transcriptional initiation site is utilized preferentially both *in vivo* and *in vitro*.

Archaeal TBP, TFB and RNA polymerase are not stably associated

In order to address the question of whether *S.shibatae* TBP and TFB function in transcriptional initiation, we raised polyclonal rabbit antisera against these proteins and against the ~100 kDa B subunit of the *S.shibatae* RNA polymerase (see Materials and methods). Importantly, Western blotting studies reveal that each antiserum recognizes its respective protein with high affinity and specificity, and does not cross-react significantly with other proteins present in *S.shibatae* cell extract (data not shown). Given that eukaryotic TBP interacts with the TFB homologue BRF, we were interested to see whether archaeal TBP, TFB and RNA polymerase might be complexed in solution. To address this issue, the anti-TBP and anti-TFB antisera were used separately to immunodeplete TBP and TFB respectively from crude *S.shibatae* transcription extracts. The resulting immunodepleted material was then tested for the presence of TBP, TFB and RNA polymerase

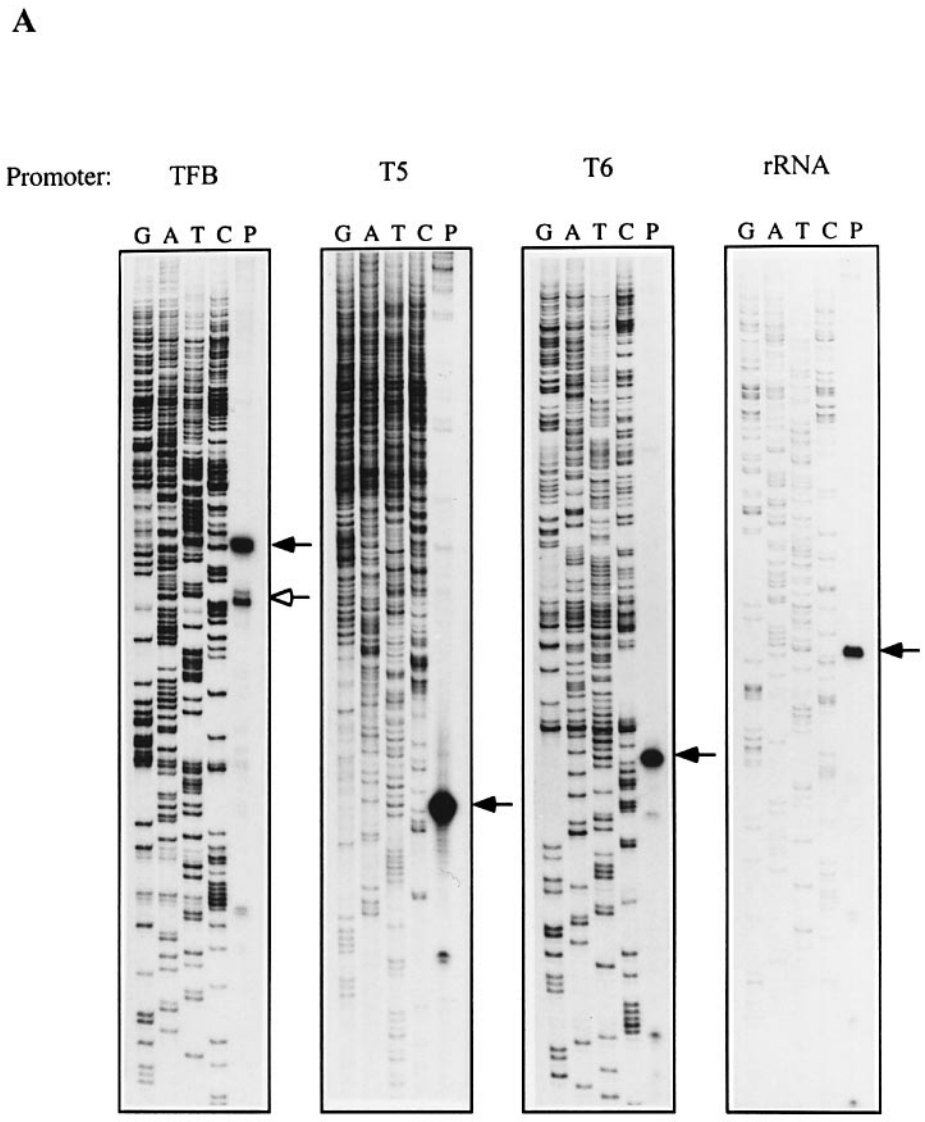

Fig. 1. Mapping of the transcription initiation sites of the *S.shibatae* TFB, T5, T6 and rRNA gene promoters. (**A**) Primer extension analysis was carried out using either T3 or T7 sequencing primers to analyse *in vitro* transcribed RNA with the *S.shibatae* extract. Each of the run-off products (P) was electrophoresed alongside its DNA sequencing reactions (G, A, T and C) also carried out with the corresponding primer. Major transcription initiation sites are designated by filled arrowheads and the transcript originating from the minor TFB promoter is shown by an open arrowhead. (**B**) DNA sequence of the four promoters. In each case the respective positions of the A-box elements and transcription initiation sites are shown. It should be noted that in each case transcription initiation sites produced *in vivo* correspond to those produced *in vitro* (see text).

by Western blotting. As shown in Figure 2A, depletion of TBP with the anti-TBP antisera does not significantly alter the amount of either TBP or the RNA polymerase in the extract. Similarly, despite the fact that the immunodepletions were conducted under conditions optimal for *in vitro* transcription, the anti-TFB antiserum removed TFB but not TBP or the RNA polymerase. These results therefore indicate that archaeal TBP, TFB and RNA polymerase do not associate appreciably with one another in solution in the absence of DNA.

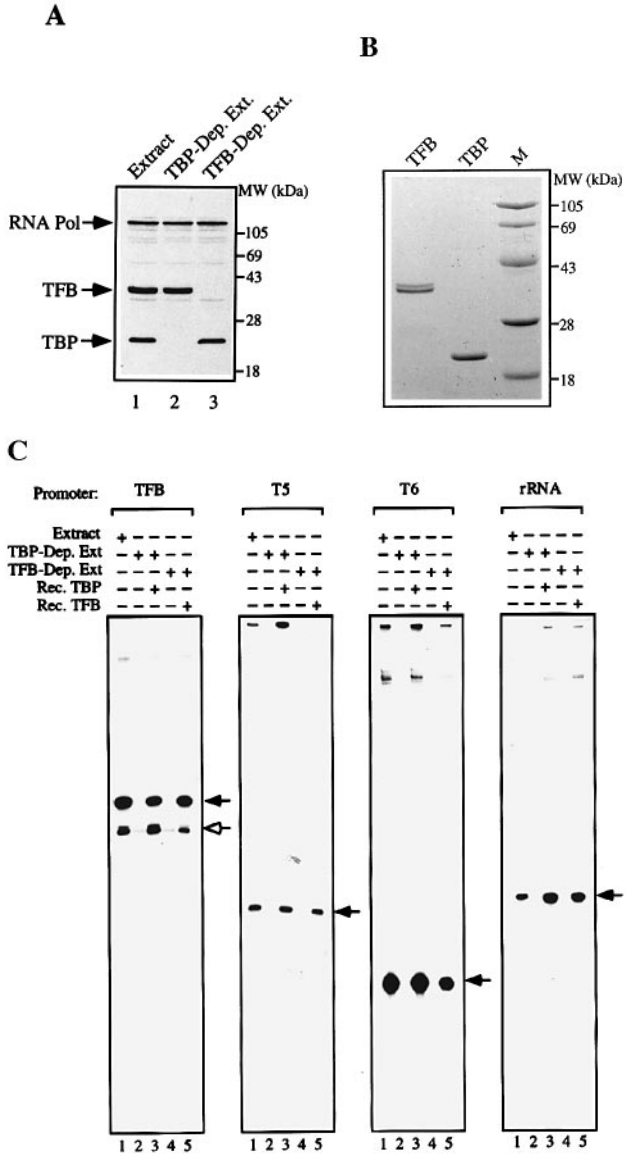

Fig. 2. Requirement of TBP and TFB for transcription from the TFB, T5, T6 and rRNA gene promoters. (**A**) Western blot analyses of the immunodepleted extracts. Total extract (lane 1), TBP and TFB-depleted extracts (lanes 2 and 3 respectively). The blot was probed simultaneously with antisera raised against *S.shibatae* TBP, TFB and rpoB. The identity of each of the proteins is shown. (**B**) Purity of the *Escherichia coli* expressed factors. N-terminally His-tagged TBP and TFB (~1 μg each) were electrophoresed alongside molecular weight markers (M) on a 12% SDS–polyacrylamide gel and detected by Coomassie staining. (**C**) *In vitro* transcription analysis of the TFB, T5, T6 and rRNA promoters using undepleted and immunodepleted extracts. Transcription reactions were carried out using undepleted (lane 1), TBP-depleted (lanes 2 and 3) and TFB-depleted (lanes 4 and 5) extract. Reactions 3 and 5 were supplemented with ~10 and 15 ng recombinant TFB and TBP respectively. The positions of the major transcripts are shown by filled arrowheads and the position of the minor transcript originating from the TFB promoter is depicted by an open arrowhead. Some sample loss occurred for lane 5 of the T6 panel, and repeats of this experiment indicated that full restoration of transcription is effected when recombinant TFB is added to the TFB-immunodepleted extract. Similarly, lane 1 of the rRNA panel is reduced by sample loss in this particular experiment.

S.shibatae TBP and TFB are required for transcription of various archaeal promoters

To assess the potential involvement of *S.shibatae* TBP and TFB in transcription, we expressed N-terminally histidine tagged derivatives of these polypeptides in *Escherichia coli*, and purified these recombinant proteins by Ni^{2+}-NTA affinity chromatography (Figure 2B). Next, we tested the TBP and TFB-depleted *S.shibatae* extracts described above for their ability to transcribe the TFB, T5, T6 and 16S/23S rRNA promoter constructs *in vitro*. In all cases, undepleted extract (12 μg; Figure 2C, lanes 1) and extract that had been mock-depleted with pre-immune sera (data not shown) gave rise to high levels of transcription. Conversely, essentially no transcription was obtained from any promoter when extracts were used that had been immunodepleted of TBP or TFB (Figure 2C, lanes 2 and 4 respectively). Most importantly, full levels of accurate transcription were restored when the TBP depleted extract was supplemented with 15 ng recombinant TBP (lanes 3), or the TFB-depleted extract was supplemented with 10 ng recombinant TFB (lanes 5). In contrast, recombinant TBP did not stimulate transcription when added to TFB depleted extracts, and TFB did not stimulate transcription with TBP depleted extracts (data not shown). As quantified by Western blotting, the amounts of recombinant TBP and TFB used in the reconstitution experiments were roughly equivalent to the amount of these proteins present in 12 μg extract. These studies therefore reveal that *S.shibatae* TBP and TFB play essential roles in transcription and that, for all promoters tested, the recombinant factors are as active as their endogenous counterparts.

Absence of TBP-associated factors in S.shibatae

In eukaryotes, TBP is found associated with distinct sets of TBP-associated factors (TAFs) in three different complexes. Immunoprecipitation of eukaryotic TBP, therefore, also leads to the removal of the TAF components of transcription factors SL-1, TFIID and TFIIIB. Although the aforementioned data clearly suggest that TBP is indispensable for transcription from all archaeal promoters, the possibility that TBP-associated factor(s) exist could not be ruled out. For instance, if there were an excess of free TBP-associated factors, these would remain in the immunodepleted extract. The subsequent addition of recombinant TBP could then lead to the assembly of TAF complexes that would participate in the formation of the preinitiation complex at the promoter. To address this issue, we employed size exclusion chromatography to compare the behaviour of endogenous TBP with that of the recombinant protein. In order to reduce non-specific interactions, the fractionation was carried out in 150 mM KCl. Furthermore, ethidium bromide was added to the *S.shibatae* crude cell extract and recombinant TBP (1.2 mg and 2 μg protein respectively) prior to fractionation to eliminate any interactions resulting from proteins binding to trace amounts of DNA. After size exclusion chromatography, the resulting fractions were analysed by Western blotting using anti-TBP antibodies. As shown in Figure 3, both recombinant TBP and endogenous TBP elute in fractions 23 and 24 from the column, consistent with a native molecular weight of 35–40 kDa. These results suggest that, as with eukaryotic TBP (Taggart and Pugh, 1996), *S.shibatae* TBP exists in solution as a homodimer. Furthermore, they provide strong support for the notion that archaeal TBP does not exist in stable complexes with other factors in solution. Similarly, we have no evidence

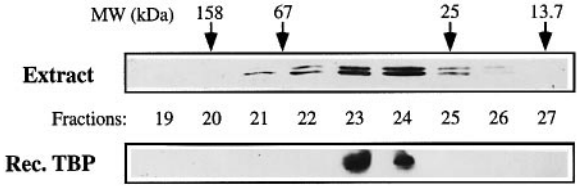

Fig. 3. Size exclusion chromatographic analysis of recombinant and endogenous TBP. Crude *S.shibatae* extract (1.2 mg) and N-terminally His-tagged TBP (2 μg) were subjected to gel filtration on a Superose-6 column and the resulting fractions were analysed by Western blotting using anti-TBP antibodies. The elution behaviour of aldolase, bovine serum albumin, chymotrypsin and RNase A (mol. wts 158, 67, 25 and 13.7 respectively) which were used as calibration markers is shown. Both endogenous and recombinant TBP elute as complexes of 35–40 kDa in fractions 23 and 24.

that TFB exists in large complexes with other factors (data not shown). Consistent with these observations, immunoaffinity purification of the endogenous TBP and TFB under low salt conditions has failed to identify any other proteins that copurify with these factors in stoichiometric amounts (data not shown).

Accurate and efficient transcription in S.shibatae requires TBP, TFB and RNA polymerase

In order to determine whether the archaeal transcription system employs basal factors in addition to TBP and TFB, it was necessary to establish a highly defined *in vitro* transcription system with purified RNA polymerase. To this end, we fractionated extracts of *S.acidocaldarius* as depicted in Figure 4A to yield an essentially homogeneous preparation of the *Sulfolobus* RNA polymerase. As revealed by tricine SDS–PAGE followed by silver staining, the RNA polymerase preparation is essentially homogeneous and, as shown previously (Zillig *et al*., 1979), comprises 10 subunits ranging in size from 10 to 122 kDa (Figure 4B). Importantly, Western blotting studies revealed that the RNA polymerase is devoid of even trace amounts of TBP or TFB (data not shown). Although *S.shibatae* and *S.acidocaldarius* are relatively closely related archaeal species, it was necessary to establish that the purified *S.acidocaldarius* enzyme is able to function effectively in the *S.shibatae* system. To do this, *S.shibatae* extract was quantitatively immunodepleted of endogenous RNA polymerase with an antibody raised against the N-terminal region of the ~100 kDa B subunit (Figure 4C). Significantly, this results in the removal of neither TBP nor TFB, which again lends support to the idea that these proteins do not form stable complexes with the RNA polymerase in solution. As shown in Figure 4D, and consistent with the existence of a single archaeal RNA polymerase, immunodepletion leads to the loss of transcription from all promoters tested (compare lanes 1 and 2). Most importantly, full levels of transcription are restored when purified *S.acidocaldarius* RNA polymerase (~15–20 ng) is added to the immunodepleted system (lanes 3 and 4). Since Western blot analyses reveal that the amount of purified RNA polymerase used in these studies is approximately equal to the amount of *S.shibatae* RNA polymerase in the undepleted extract, we conclude that the purified *S.acidocaldarius* RNA polymerase has high specific activity and is fully able to function in the *S.shibatae* system.

Next, we used the purified RNA polymerase, both alone and in the presence of recombinant TBP and/or TFB, in *in vitro* transcription studies (Figure 5). Significantly, the minor TFB promoter is transcribed efficiently by the purified RNA polymerase alone (left panel; lane 1) and is not stimulated significantly by the addition of TBP and/or TFB (lanes 2–4). This indicates that the region spanning the minor transcription initiation site of the TFB gene has an intrinsic capacity to recruit the RNA polymerase in the absence of additional basal factors. All other promoters tested, however, are not transcribed by the RNA polymerase alone (lanes 1) and, furthermore, no stimulation of transcription is obtained when the RNA polymerase is supplemented with either TBP or TFB alone (Figure 5, lanes 2 and 3 respectively). In contrast, when the RNA polymerase is supplemented with both TBP and TFB, the major upstream TFB promoter as well as the T5, T6 and 16/23S rRNA promoters are all transcribed very efficiently (lanes 4). Although the data in Figure 5 were obtained with supercoiled templates, similar high levels of transcription are also observed when RNA polymerase, TBP and TFB are used to transcribe linearized templates (data not shown). These results indicate that *S.shibatae* RNA polymerase, TBP and TFB are sufficient to direct accurate transcription and that, with the exception of the minor TFB promoter, transcription is abrogated when any one of these factors is omitted.

To see whether full levels of transcription are restored by the three archaeal factors, we performed transcription assays in parallel with crude extract and with the purified components (Figure 6). As estimated by quantitative Western blotting, the amounts of purified factors employed in this experiment were equal to the amounts of these factors present in the crude extract. Importantly, with the exception of the TFB upstream promoter, recombinant TBP and TFB together with purified RNA polymerase yield at least as much transcription as the crude extract. Moreover, as with the crude system, the T6 and T5 promoters produce the highest levels of transcription in the defined system. These results therefore imply that, at least for the promoters used, no additional factors exist in crude *S.shibatae* extracts that are able to enhance transcription directed by RNA polymerase, TBP and TFB. Furthermore, they suggest strongly that the relative strengths of archaeal promoters are defined primarily by their interactions with the basal transcription apparatus.

Nevertheless, it is noteworthy that the upstream TFB promoter is transcribed less efficiently in the defined system than in the crude extract, possibly reflecting the existence of a gene-specific transcriptional activator. In addition, it is clear that the defined system produces higher levels of transcription from the rRNA promoter than is apparent in the crude system. One explanation for this is that a factor exists in the crude *S.shibatae* extract that negatively regulates the 16S/23S rRNA promoter. Consistent with this, previous work has shown that a negative regulatory region exists ~90–200 bp upstream of the rRNA promoter (Reiter *et al.*, 1990).

Discussion

In this manuscript, we have shown that *S.shibatae* RNA polymerase, TBP and TFB are sufficient to direct accurate

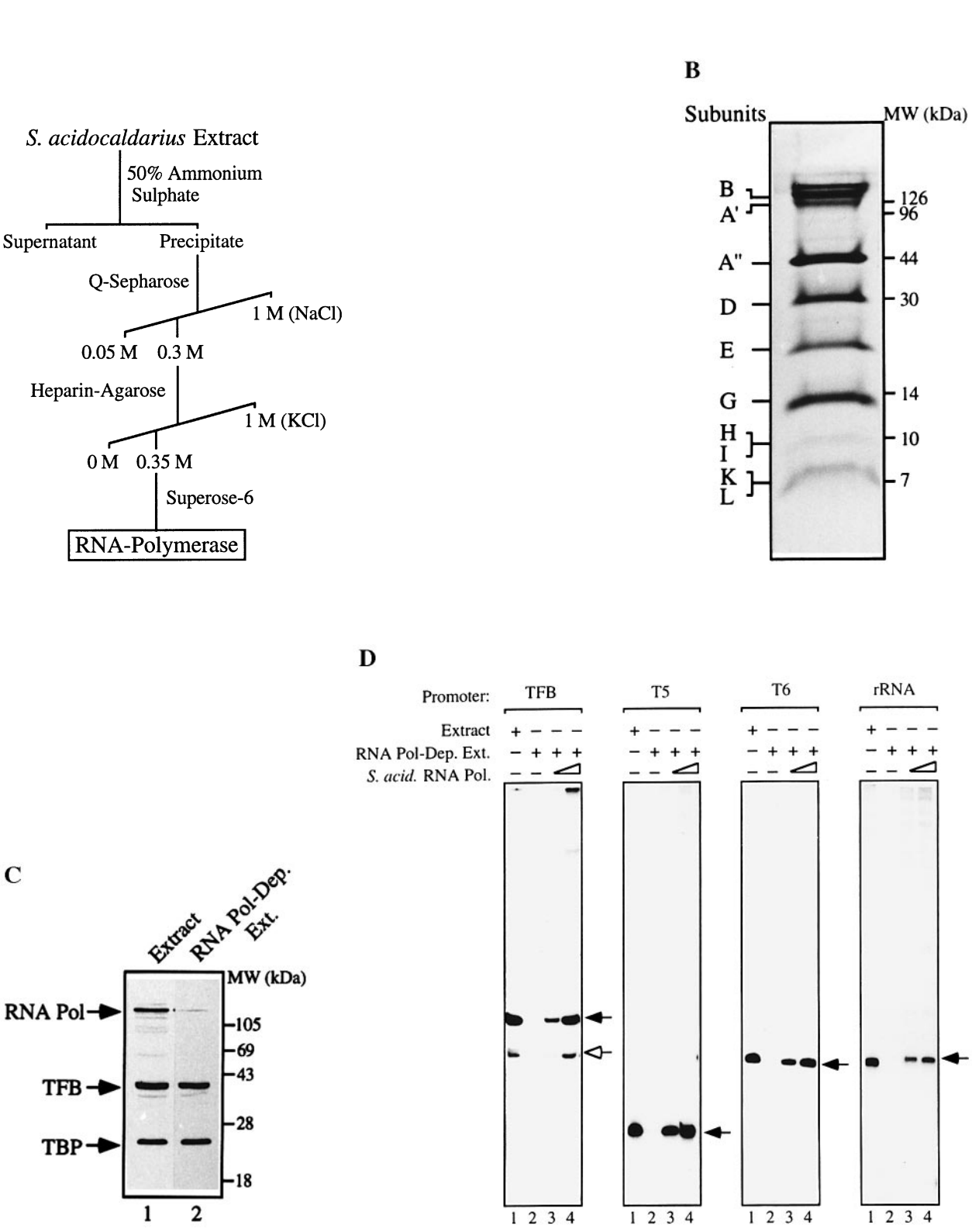

Fig. 4. Purified *S.acidocaldarius* RNA polymerase is active in the *S.shibatae* system. (**A**) Outline of the scheme used for the purification of the *S.acidocaldarius* RNA polymerase. (**B**) Purity of the RNA polymerase was established by electrophoresing ~1 µg protein on a 18% tricine gel system prior to silver staining. The identity of the 10 RNA polymerase subunits is shown. (**C**) Western blot analysis of the RNA polymerase immunodepleted extract. Total extract (lane 1) and RNA polymerase depleted extract (lane 2). The blot was probed simultaneously with antisera raised against *S.shibatae* TBP, TFB and rpoB. The position of each of the proteins is indicated. (**D**) *In vitro* transcription analysis of the four promoters using undepleted and RNA polymerase depleted extracts. Transcription reactions were carried out using *S.shibatae* undepleted (lane 1) and RNA polymerase depleted (lanes 2, 3 and 4) extract. Reactions 3 and 4 were supplemented with ~3 and 15 ng *S.acidocaldarius* RNA polymerase respectively. The positions of the major transcripts are shown by filled arrowheads whereas the open arrowhead identifies the transcript which originates from the downstream TFB promoter.

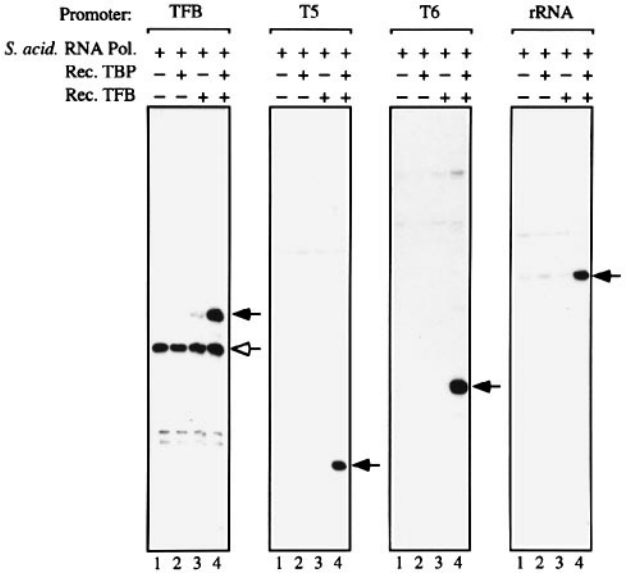

Fig. 5. Transcription using defined components. *In vitro* transcription analyses of the TFB, T5, T6 and rRNA gene promoters were carried using *S.acidocaldarius* RNA polymerase (~15 ng). Reactions 2 and 4 were supplemented with recombinant TBP (15 ng), while reactions 3 and 4 received 10 ng recombinant TFB. The positions of the major transcripts are shown by filled arrowheads and the position of the minor transcript originating from the downstream TFB promoter is shown by an open arrowhead.

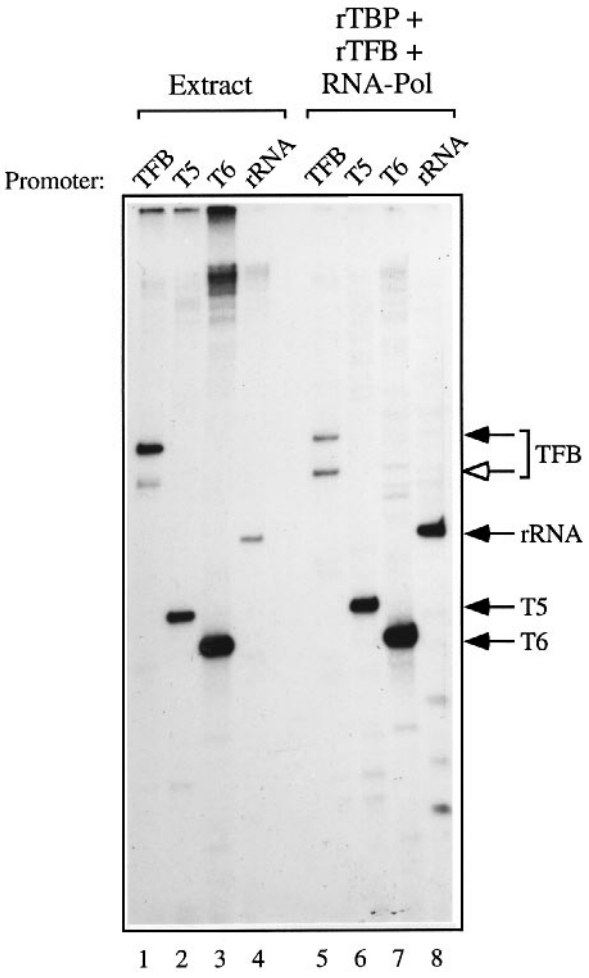

Fig. 6. Monitoring the efficiency of the two transcription systems. *In vitro* transcription assays were carried out in parallel on the TFB, T5, T6 and rRNA gene promoters using the crude *S.shibatae* extract (12 μg) and the defined system containing equivalent amounts of *S.acidocaldarius* RNA polymerase (~15 ng), recombinant TBP and TFB (15 and 10 ng respectively). The position of each of the major transcripts is shown by filled arrowheads whereas the position of the minor transcript originating from the downstream TFB promoter is indicated by an open arrowhead.

transcription from a variety of archaeal promoters, and that transcription is abrogated when any one of these factors is omitted (see below for the exception of the downstream TFB promoter). These findings are consistent with recent studies in the *Pyrococcus furiosus* system, where partially purified factors were employed to study transcription initiation from the glutamate dehydrogenase gene promoter (Hethke *et al*., 1996). Although it remains a formal possibility that transcription is stimulated further by additional basal factors that are present in the crude *S.shibatae* extract, several lines of evidence indicate that this is not the case. First, when TBP or TFB is removed from extracts by immunodepletion, full levels of transcription are recovered when the system is complemented by recombinant TBP or TFB respectively. Second, gel filtration analyses show that the respective elution profiles of recombinant and endogenous archaeal TBP and TFB are essentially identical. Third, no clear homologues of any additional eukaryotic-type basal transcription factors, including TAFs, are evident in the recently published complete genomic DNA sequence of the archaeon *Methanococcus jannaschii* (Bult *et al.*, 1996). Fourth, using amounts of TBP, TFB and RNA polymerase that are equivalent to those found in the crude *S.shibatae* cell lysate, the defined system directs transcription as efficiently as crude *S.shibatae* cell extracts. Finally, we have been unable to detect additional stimulatory activities through fractionation of the *S.shibatae* extract (data not shown).

As is the case in eukaryotes, we have shown that archaeal TBP, TFB and RNA polymerase do not interact with one another stably in the absence of DNA, and appear to associate only in the context of promoter DNA. We have also shown that archaeal TBP is like its eukaryotic counterpart in that it appears to exist in solution as a homodimer (Taggart and Pugh, 1996). Furthermore, as is the case for eukaryotic TBP and TFIIB, archaeal TBP alone only binds to the A-box sequence weakly and requires TFB in order to interact efficiently and stably with the DNA template (Rowlands *et al*., 1994; Qureshi *et al*., 1995b). A further parallel between the archaeal and eukaryotic systems is that, under most circumstances, the RNA polymerase is only recruited subsequent to the establishment of a ternary complex containing the promoter, TBP and TFB (TFIIB). Although a large number of proteins constitute the general transcriptional apparatus of eukaryotes, it has been shown that minimal transcription of certain supercoiled class II templates only requires RNA polymerase II, TBP and TFIIB (Parvin and Sharp, 1993). The minimal eukaryotic system therefore bears a strong resemblance to the archaeal system, which we have shown employs RNA polymerase, TBP and TFB. Significantly, we find that this set of archaeal factors can transcribe negatively supercoiled templates and linearized templates with similar efficiency (unpublished data). This therefore contrasts with the situation in eukaryotes, where

additional basal factors are usually required, at least in part, to facilitate melting around the initiation site and to allow promoter escape by the RNA polymerase (Goodrich and Tjian, 1994). One attractive possibility is that additional activities are not required in the *Sulfolobus* system because promoter opening is already facilitated by the fact that this organism grows at temperatures in the range 70–80°C.

One of the most significant conclusions resulting from our work is that the relative strengths of various archaeal promoters are the same in the reconstituted *Sulfolobus* system as they are in crude unfractionated extracts. This suggests that promoter strength in Archaea is defined primarily by interactions with the basal transcriptional apparatus. One exception to this general conclusion is the 16S/23S rRNA promoter, which is transcribed more efficiently in the defined system than in the crude system. This possibly reflects the existence of a negative regulatory factor in crude extracts, an idea that is consistent with previous studies, which have defined a negative upstream element (Reiter *et al*., 1990). In addition, we find that the reconstituted *in vitro* system transcribes the upstream and downstream TFB promoters with equal efficiency, whereas transcripts originate predominantly from the upstream promoter *in vivo*. Although the reason for this difference is currently unknown, it is tempting to speculate that the efficiency of transcription of one or other promoter is regulated *in vivo* to control TFB levels. In this regard, it is noteworthy that we have recently identified an activity in *Sulfolobus* extracts that interacts specifically with the initiator region of the minor TFB promoter (S.A.Qureshi and S.P.Jackson, unpublished data), suggesting that this factor might positively or negatively regulate the usage of this transcription start site.

Although it has been reported previously that the *Sulfolobus* RNA polymerase has an intrinsic ability to initiate transcription from the 16S/23S rRNA promoter (Hüdepohl *et al.*, 1990), these studies employed a large molar excess of the purified enzyme. Indeed, our data reveal that, when present at physiological levels, the RNA polymerase alone is unable to initiate transcription accurately from the rRNA promoter and is also unable to transcribe the T5, T6 and upstream TFB promoters. In marked contrast, however, we find that the downstream TFB promoter is transcribed accurately by low amounts of purified RNA polymerase in the absence of any other factors. Notably, however, this promoter is not transcribed by either TBP or TFB depleted crude cell extracts. This suggests that negatively regulating factors, which may be competed out by TBP or TFB, might interfere with the binding of RNA polymerase to DNA in the unfractionated system. It will clearly be of great interest to define the mechanism of RNA polymerase recruitment to the downstream TFB promoter and, in particular, to define which subunits of the archaeal polymerase mediate promoter recognition. Although most archaeal promoters appear to possess A-box elements and are therefore likely to employ TBP for their transcription, it is intriguing to note that the SSV $T_{induced}$ promoter, which is activated after cells have been irradiated with UV light, lacks an A-box (Reiter *et al.*, 1988). It will therefore be of interest to ascertain whether the archaeal RNA polymerase has an inherent ability to recognize this promoter and whether this promoter can be transcribed in the absence of TBP. Given the strong regulation of the SSV $T_{induced}$ promoter, it is tempting to speculate that the lack of an A-box motif may reduce the basal activity of certain archaeal promoters and, thus, provide greater opportunities for regulatory control.

In conclusion, our studies provide strong support for the notion that the transcriptional apparatus of Archaea is fundamentally similar to the nuclear transcription systems of eukaryotes and is significantly different from that of Bacteria. From an evolutionary point of view, this suggests that a sophisticated transcription system containing TBP, TFB and a complex RNA polymerase was already in place in the last common ancestor of Archaea and Eucarya, before these two lineages diverged. Furthermore, it suggests that the additional basal factors found in eukaryotic systems evolved subsequent to this divergence, presumably to provide additional scope for regulatory mechanisms and to allow transcription to be linked to other nuclear events such as DNA repair, RNA processing and RNA export.

Materials and methods

Plasmid construction

All DNA modifying enzymes were purchased from Boehringer Mannheim and were used according to manufacturer's instructions. Sequences of oligonucleotides employed in these studies read 5′ to 3′. The T5, T6 and TFB promoter regions were amplified by PCR using the following sets of oligonucleotides: GATCGAATTCGAGTTTGTGCCATATTCCCAT and GATCAAGCTTCTTCACTCTCTCTTCTAGCTC for T5; GATCGAATTCGATTGGAATCGAAACGGTCAC and GATCAAGCTTCTTCTTCTTTTCGGGTTTGGG for T6; GATCAAGCTTAACGCTTAATTTATACTG and GATCAAGCTTCCCCCCTCTCCGCATCGAATA for TFB. PCR was performed under standard conditions and the products purified on PCR purification columns (Qiagen). The TFB PCR product was digested with *Hin*dIII and was cloned into the corresponding site of pBluescript (Stratagene). All other PCR products were digested with *Eco*RI and *Hin*dIII and were cloned into *Eco*RI/*Hin*dIII digested pBluescript II SK (Stratagene). The 16S/23S rRNA promoter region was cloned by excising the 219 bp *Dra*I–*Hin*dIII (–127 to +92) fragment and ligating it into *Eco*RV/*Hin*dIII digested pBluescript. For cloning of the TFB open reading frame, the following oligonucleotides were used in a standard PCR: GGATGGATCCGTGTTGTATTTGTCTGAAGAAAAT and GTAGAAGCTTATGCGGAATCTAACATTCTAGTTA. The resulting fragment was purified and digested with *Bam*HI and *Hin*dIII and was cloned into the corresponding sites of pQE30 (Qiagen). The plasmid pQErpoB was constructed by using the oligonucleotides GATCGGATCCAAAGATCCGGGTGGTTACTT and GATCAAGCTTAGCACTAGCTAAATAATATCCTTTCTT, which were used to generate by PCR a DNA fragment corresponding to amino acid residues 123–296 of the B subunit of *S.shibatae* RNA polymerase. This was then cloned into the *Bam*HI and *Hin*dIII sites of pQE30. All plasmids were purified by CsCl density gradient centrifugation and were sequenced with Sequenase version 2.0 (US Biochemicals).

Expression of proteins and preparation of antisera

TBP was cloned, expressed and purified as described previously (Qureshi *et al*., 1995a). For expression of TFB and a region of rpoB (amino acids 123–296) the respective plasmids were used to transform M15[pREP4] cells. The proteins were expressed and purified according to the instructions provided in the QIAexpressionist booklet (Qiagen). Antiserum against rpoB was generated by injecting 100 μg of the 22 kDa histidine-tagged protein into a rabbit at monthly intervals.

Preparation of S.shibatae whole cell extracts, immunodepletion of extracts and Western blotting

Whole cell extracts were prepared according to Hüdepohl *et al*. (1990). TFB, TBP and RNA polymerase-depleted extracts were prepared by taking 25 μl protein A–Sepharose beads (Pharmacia) and washing them with phosphate-buffered saline (PBS). Subsequently, 25 μl antisera

raised against bacterially expressed *P.woesei* TFB, *S.shibatae* TBP and *S.shibatae* RNA polymerase subunit B was added to the beads and left on ice for 1 h with occasional mixing. After washing with PBS, the beads were incubated with 150 µl (~1.8 mg) extract and the mixture was left on ice for 3 h with occasional mixing. The beads were then removed by centrifugation and the depleted extract collected. For Western blotting, proteins were transferred from SDS–polyacrylamide gels to nitrocellulose membranes, and these were probed with the appropriate antibodies (each diluted 1:2000) according to standard procedures. Detection of bound antibodies was accomplished using the ECL system (Amersham).

Purification of S.acidocaldarius RNA polymerase

Twenty-five grammes wet cell weight of *S.acidocaldarius* DSM 639 were resuspended in 100 ml 50 mM Tris–HCl pH 7.8, 100 mM NaCl, 10 mM β-mercaptoethanol and lysed by three passes through a French press (Gaulin). The extract was clarified by centrifugation at 20 000 r.p.m. in an SS-34 rotor (Sorvall). Solid ammonium sulfate (Sigma) was added to 50% saturation at 4°C and the precipitate was collected by centrifugation at 20 000 r.p.m. in an SS-34 rotor. The pellet was resuspended in 50 ml buffer A (50 mM Tris–HCl pH 7.8, 10 mM β-mercaptoethanol and 10% glycerol) containing 50 mM NaCl and dialysed against 100 vol of this buffer overnight. Following dialysis the solution was centrifuged to remove any precipitates and the supernatant was applied to a 15 ml Q-Sepharose (Pharmacia) column at a flow rate of one column volume per hour. The column was developed with a 150 ml gradient, from 50 mM to 1 M NaCl, in buffer A. Based on Western blotting, using the *S.shibatae* anti-rpoB antibody, and non-specific transcription initiation assays the RNA polymerase was found to elute at ~300 mM NaCl. The positive fractions were pooled, dialysed against 100 vol buffer A and applied to a 2.5 ml heparin–agarose (Sigma) column which was subsequently developed with a 10 column volume gradient, from 0 to 1 M KCl in buffer A. RNA polymerase eluted at 350 mM KCl. Positive fractions were concentrated by spin dialysis through Centricon C100 concentrators (Amicon) and applied to a Superose-6 column (Pharmacia) equilibrated with buffer A + 300 mM KCl. The RNA polymerase eluted from this column with an apparent molecular weight of 410 kDa, in close agreement with the predicted molecular mass derived from the combined molecular weights of its various subunits (Zillig *et al*., 1979; Klenk *et al*., 1992).

In vitro transcription and primer extension analyses

In vitro transcription reactions and primer extension analyses were carried out using end-labelled T3 (CGAAATTAACCCTCACTAAAGGGAAC) and T7 (TACGACTCACTATAGGGCGAATTGGG) primers as described previously (Qureshi *et al*., 1995a), with the exception that transcription buffer was supplemented with 80 mM KCl. All primer extension analyses were carried out using a large molar excess of either the T3 or T7 oligonucleotide.

Acknowledgements

We thank members of the S.P.Jackson laboratory for their encouragement. This work was funded by grant G9420496MB from the Medical Research Council (UK) and by Grant F/770 from the Leverhulme Trust.

References

Barns,S.M., Delwiche,C.F., Palmer,J.D. and Pace,N.R. (1996) Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proc. Natl Acad. Sci. USA*, **93**, 9188–9193.

Baumann,P., Qureshi,S.A. and Jackson,S.P. (1995) Transcription: new insights from studies on Archaea. *Trends Genet.*, **11**, 279–283.

Berghofer,B., Krockel,L., Kortner,C., Truss,M., Schallenberg,J. and Klein,A. (1988) Relatedness of archaebacterial RNA polymerase core subunits to their eubacterial and eukaryotic equivalents. *Nucleic Acids Res.*, **16**, 8113–8128.

Bult,C.J. *et al.* (1996) Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science*, **273**, 1058–1073.

Creti,R., Londei,P. and Cammarano,P. (1993) Complete nucleotide sequence of an archaeal (*Pyrococcus woesei*) gene encoding a homolog of eukaryotic transcription factor IIB (TFIIB). *Nucleic Acids Res.*, **21**, 2942.

DeLong,E.F., Wu,K.Y., Prezelin,B.B. and Jovine,R.V.M. (1994) High abundance of Archaea in Antarctic marine picoplankton. *Nature*, **371**, 695–697.

Eick,D., Wedel,A. and Heumann,H. (1994) From initiation to elongation: comparison of transcription by prokaryotic and eukaryotic RNA polymerases. *Trends Genet.*, **10**, 292–296.

Fox,G.E., Stackebrandt,E., Hespell,R.B., Gibson,J., Maniloff,J., Dyer,T.A., Wolfe,R.S., Balch,W.E., Tanner,R.S., Magrum,L. J., Zablen,L.B., Blakemore,R., Gupta,R., Bonen,L., Lewis,B.J., Stahl,D.A., Luehrsen,K.R., Chen,K.N. and Woese,C.R. (1980) The phylogeny of prokaryotes. *Science*, **209**, 457–463.

Frey,G., Thomm,M., Brudigam,B., Gohl,H.P. and Hausner,W. (1990) An archaebacterial cell-free transcription system. The expression of tRNA genes from *Methanococcus vannielii* is mediated by a transcription factor. *Nucleic Acids Res.*, **18**, 1361–1367.

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

Hausner,W. and Thomm,M. (1993) Purification and characterization of a general transcription factor, aTFB, from the Archaeon *Methanococcus thermolithotrophicus*. *J. Biol. Chem.*, **268**, 24047–24052.

Hausner,W., Frey,G. and 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.

Hempel,W.M., Cavanaugh,A.H., Hannan,R.D., Taylor,L. and Rothblum,L.I. (1996) The species-specific RNA polymerase I transcription factor SL-1 binds to upstream binding factor. *Mol. Cell. Biol.*, **16**, 557–563.

Hernandez,N. (1993) TBP, a universal eukaryotic transcription factor? *Genes Dev.*, **7**, 1291–1308.

Hethke,C., Geerling,A.C., Hausner,W., de Vos,W.M. and Thomm,M. (1996) A cell-free transcription system for the hyperthermophilic archaeon *Pyrococcus furiosus*. *Nucleic Acids Res.*, **24**, 2369–2376.

Hüdepohl,U., Reiter,W.-D. and Zillig,W. (1990) *In vitro* transcription of two rRNA genes of the archaebacterium Sulfolobus sp. B12 indicates a factor requirement for specific initiation. *Proc. Natl Acad. Sci. USA*, **87**, 5851–5855.

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

Khoo,B., Brophy,B. and Jackson,S.P. (1994) Conserved functional domains of the RNA polymerase III general transcription factor BRF. *Genes Dev.*, **8**, 2879–2890.

Klenk,H.-P., Palm,P., Lottspeich,F. and Zillig,W. (1992) Component H of the DNA-dependent RNA polymerases of Archaea is homologous to a subunit shared by the three eucaryal nuclear RNA polymerases. *Proc. Natl Acad. Sci. USA*, **89**, 407–410.

Langer,D., Hain,J., Thuriaux,P. and Zillig,W. (1995) Transcription in Archaea: similarity to that in Eucarya. *Proc. Natl Acad. Sci. USA*, **92**, 5768–5772.

Marsh,T.L., Reich,C.I., Whitelock,R.B. and Olsen,G.J. (1994) Transcription factor IID in the Archaea: sequences in the *Thermococcus celer* genome would encode a product closely related to the TATA-binding protein of eukaryotes. *Proc. Natl Acad. Sci. USA*, **91**, 4180–4184.

Orphanides,G., Lagrange,T. and Reinberg,D. (1996) The general transcription factors of RNA polymerase II. *Genes Dev.*, **10**, 2657–2683.

Ouzounis,C. and Sander,C. (1992) TFIIB, an evolutionary link between the transcription machineries of archaebacteria and eukaryotes. *Cell*, **71**, 189–190.

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

Puhler,G., Leffers,H., Gropp,F., Palm,P.K., Klenk,H.-P., Lottspeich,F., Garrett,R.A. and Zillig,W. (1989) Archaebacterial DNA-dependent RNA polymerases testify to the evolution of the eukaryotic nuclear genome. *Proc. Natl Acad. Sci. USA*, **86**, 4569–4573.

Qureshi,S.A. and Jackson,S.P. (1997) Relatedness between various transcription systems. In Bradshaw,R.A. and Purton,M. (eds), *Modular Texts in Molecular and Cell Biology.* Chapman and Hall, London, UK, in press.

Qureshi,S.A., Baumann,P., Rowlands,T., Khoo,B. and Jackson,S.P. (1995a) Cloning and functional analysis of the TATA binding protein from *Sulfolobus shibatae*. *Nucleic Acids Res.*, **23**, 1775–1781.

Qureshi,S.A., Khoo,B., Baumann,P. and Jackson,S.P. (1995b) Molecular cloning of the transcription factor TFIIB homolog from *Sulfolobus shibatae*. *Proc. Natl Acad. Sci. USA*, **92**, 6077–6081.

Reiter,W.-D., Palm,P., Voos,W., Kaniecki,J., Grampp,B., Schulz,W. and Zillig,W. (1987) Putative promoter elements for the ribosomal RNA genes of the thermoacidophilic archaebacterium *Sulfolobus* sp. strain B12. *Nucleic Acids Res.*, **15**, 5581–5595.

Reiter,W.-D., Palm,P. and Zillig,W. (1988) Analysis of transcription in the archaebacterium *Sulfolobus* indicates that archaebacterial promoters are homologous to eukaryotic pol II promoters. *Nucleic Acids Res.*, **16**, 1–19.

Reiter,W.-D., Hudepohl,W. and Zillig,W. (1990) Mutational analysis of an archaebacterial promoter: Essential role of a TATA box for transcription efficiency and start-site selection *in vitro*. *Proc. Natl Acad. Sci. USA*, **87**, 9509–9513.

Roeder,R.G. (1996) The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem. Sci.*, **21**, 327–335.

Rowlands,T., Baumann,P. and Jackson,S.P. (1994) The TATA-binding protein: a general transcription factor present in both eukaryotes and archaebacteria. *Science*, **264**, 1326–1329.

Schleper,C., Kubo,K. and Zillig,W. (1992) The particle SSV1 from the extremely thermophilic archaeon *Sulfolobus* is a virus: Demonstration of infectivity and of transfection with viral DNA. *Proc. Natl Acad. Sci. USA*, **89**, 7645–7649.

Taggart,A.K.P. and Pugh,B.F. (1996) Dimerization of TFIID when not bound to DNA. *Science*, **272**, 1331–1333.

Tanese,N., Pugh,B.F. and Tjian,R. (1991) Coactivators for a proline-rich activator purified from the multisubunit human TFIID complex. *Genes Dev.*, **5**, 2212–2224.

Verrijzer,C.P. and Tjian,R. (1996) TAFs mediate transcriptional activation and promoter selectivity. *Trends Biochem. Sci.*, **21**, 338–342.

Werner,M., Chaussivert,N., Willis,I.M. and Sentenac,A. (1993) Interaction between a complex of RNA polymerase III subunits and the 70 kDa component of transcription factor IIIB. *J. Biol. Chem.*, **268**, 20721–20724.

White,R.J. (1994) *RNA Polymerase III Transcription*. R.G.Landes Co, Austin, TX.

White,R.J. and Jackson,S.P. (1992) The TATA-binding protein: a central role in transcription by RNA polymerases I, II and III. *Trends Genet.*, **8**, 284–288.

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

Woese,C.R., Kandler,O. and Wheelis,M.L. (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria and Eucarya. *Proc. Natl Acad. Sci. USA*, **87**, 4576–4579.

Zillig,W., Palm,P., Klenk,H.-P., Langer,D., Hudepohl,U., Hain,J., Lanzendorfer,M. and Holz,I. (1993). Transcription in Archaea. In Kates,M., Kushner,D. J. and Matheson,A.T. (eds), *The Biochemistry of Archaea (Archaebacteria)*. Elsevier, Amsterdam, The Netherlands, pp. 367–391.

Zillig,W., Stetter,K.O. and Janekovic,D. (1979) DNA-dependent RNA polymerase from the Archaebacterium *Sulfolobus acidocaldarius*. *Eur. J. Biochem.*, **96**, 597–604.

Received on December 6, 1996; revised on January 14, 1997