Activation of a dual adenovirus promoter containing nonconsensus TATA motifs in *Schizosaccharomyces pombe*: role of TATA sequences in the efficiency of transcription

Sathyamangalam Swaminathan, Pawan Malhotra, Chitra F.Manohar⁺, Ravi Dhar¹ and Bayar Thimmapaya*

Robert H.Lurie Cancer Center and Microbiology and Immunology Department, Northwestern University Medical School, 303 East Chicago Avenue, Chicago, IL 60611 and ¹Laboratory of Molecular Virology, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA

Received December 14, 1992; Revised and Accepted April 27, 1993

ABSTRACT

The role of TATA elements in the expression of a mammalian promoter was investigated in the fission yeast *Schizosaccharomyces pombe*, by studying the human adenovirus E2-early promoter. This is a unique dual promoter with two nonconsensus TATA elements directing transcription from two cap sites, +1 and –26. A sequence TTAAGA provides the TATA box function for the +1 promoter, whereas a sequence TAAATT, with a closer resemblance to the consensus (TATAA/TA) provides this function for the –26 promoter. Yet, in human cells, the +1 promoter is transcribed about 20 fold more efficiently than the –26 promoter. We found that both promoters are transcribed faithfully in *S.pombe* with start sites identical or close to those found in human cells. Surprisingly, the relative ratio of expression for the +1 and –26 promoters was exactly reversed in *S.pombe* cells. This reversal appeared to be due to the relatively weak binding of *S.pombe* TATA binding protein to the TTAAGA motif, rather than to its rate of dissociation. Furthermore, we show that in *S.pombe*, promoter expression correlates well with the nucleotide sequence of the TATA element rather than the context in which it is placed. By contrast, it is the context of the TATA element, rather than its nucleotide sequence that appears to be critical for promoter expression in human cells. Our data suggest the existence of one or more additional factors in human cells that permit the utilization of nonconsensus TATA elements. *S.pombe* appears to lack these factors.

INTRODUCTION

The molecular mechanism of transcriptional activation of RNA polymerase II genes is apparently highly conserved across a broad

spectrum of eucaryotic organisms ranging from humans to yeast (reviewed in 1,2). Both structural and functional similarities between mammalian and yeast transcriptional activators have been found. The most striking example is the TATA box binding protein (TBP) which shows a high degree of conservation in its carboxy terminal 180 amino acid residues between different organisms (3). TBP appears to be interchangeable between yeast and human *in vitro* transcription systems (reviewed in 4). TBP binds to the TATA element, found in most eucaryotic polymerase II promoters about 30 basepairs (bp) upstream of the RNA start site and plays a pivotal role in determining the rate of transcription (reviewed in 2–4). In *Saccharomyces cerevisiae*, however, TATA boxes are usually found further upstream (2,5).

The conserved region of TBP recognizes the TATA element (6,7, see ref.3 for a review) apparently through contacts in the minor groove of the DNA helix (8,9). This binding which proceeds in a two-step pathway (10) represents an early event in core promoter activation, triggering an ordered stepwise assembly of other general transcription factors and RNA polymerase II (reviewed in 3,4). The mammalian TBP in the native form is tightly associated with several additional polypeptides (TBP associated factors or TAFs) which apparently can mediate productive interaction with promoter specific activators (11–13). Recently TBP has been shown to be essential for transcription of all three classes of eucaryotic genes; its association with TAFs appears to specify which polymerase can initiate transcription at a given gene (reviewed in 14).

Although the majority of polymerase II promoters contain the TATA sequence element, a significant number of promoters contain sequence elements that deviate from the consensus sequence (TATAA/TA), but presumably serve the role of TATA motifs. Promoters with both consensus and nonconsensus TATA motifs are transcribed efficiently in mammalian cells. The mechanism of interaction of these non-TATA type sequence

* To whom correspondence should be addressed

⁺ Present address: Department of Medicine, Northwestern University Medical School, 303 E. Chicago Ave, Chicago, IL 60611, USA

motifs with TBP and their role in transcription initiation is not well understood. Adenovirus (Ad) E2-early promoter is one of the viral early promoters that lacks consensus TATA motifs (15). The structure of this promoter is complex in that it contains two overlapping transcriptional control regions that direct transcription from the +1 and -26 positions (15-18). Both promoters lack the consensus TATA sequence, but have in its place sequence motifs TTAAGA and TAAATT, centered approximately 30 bp upstream of the +1 and -26 cap sites, respectively (15-18). *In vitro* and *in vivo* studies have shown that the two start sites of this dual promoter are indeed controlled by these two TATA-like sequences (15,18). Recently, the TBP gene has been cloned from a number of different organisms, ranging from yeast to humans (19-30). In all these instances, only one TBP gene has been identified with the exception of maize (30) and *Arabidopsis* (26) in which two genes that encode TBP have been found. It is therefore very likely that a single TBP binds to both consensus and nonconsensus TATA motifs.

Very little information currently exists on *S.pombe* promoters and in particular, the range of DNA sequences that can serve as functional TATA elements in this organism. A very extensive study of this question has been carried out in *S.cerevisiae* (31-33). A purified TFIID protein preparation from *S.cerevisiae* was shown to bind to a variety of nonconsensus DNA sequences with high affinity (32). However, Harbury and Struhl showed that almost all single base-pair substitutions of the TATAAA element of the *his 3* promoter abolished the ability to activate transcript *in vivo* indicating that in the *in vivo* assays, yeast cells are very sensitive to alterations in the TATA element and capable of discriminating sequences that deviate from the consensus TATA sequence (33). It is becoming increasingly apparent that the fission yeast, *S.pombe*, rather than *S.cerevisiae* is closer in many respects to higher eucaryotes (34-39). There are indications that these two yeasts have diverged during evolution from each other as much as they have from humans (34,40). It has been suggested that *S.pombe* may be a better model system for genetic studies of transcription factor mechanisms (35,41). The TBPs of *S.cerevisiae* and *S.pombe* show an exceptionally high degree of homology (93% identity; refs 20,21) and yet differ significantly with respect to transcription initiation: In *S.pombe*, transcription typically initiates at a distance of 25-47 nucleotides downstream of the TATA element (37-39,42), while in *S.cerevisiae* the spacing requirements are nonstringent with the distance between the TATA element and the initiation site ranging from 40 to 120 nucleotides (2,5). Further, in *S.cerevisiae* the precise site of initiation appears to be determined not by the TATA element, but instead by an initiator element (2,5). For these reasons, we are interested in comparing the transcriptional regulation mechanisms between *S.pombe* and human cells, using our E2-early promoter model. As *S.pombe* is equally amenable to genetic analysis as *S.cerevisiae* and is closer than the latter to higher eucaryotes, particularly with respect to the transcriptional machinery, it has the potential to serve as a useful system for investigating higher eucaryotic transcription.

An intriguing feature of the E2-early promoter is that in human cells the +1 cap site is utilized about 20 fold more efficiently than the -26 cap site even though the TATA element, TAAATT, of the -26 promoter bears a closer resemblance to the consensus TATA sequence compared to TTAAGA sequence of the +1 promoter (15-18). As a first step towards understanding this phenomenon and developing *S.pombe* as a genetic system for analysing mammalian transcription, we investigated whether the

expression of the E2-early promoter in *S.pombe* would display the same 'anomalous' preference for transcription at the +1 cap site as in human cells. Surprisingly, we found that the relative efficiency with which the two promoters were expressed in *S.pombe* cells was dramatically altered, with the -26 promoter functioning as the major promoter. This reversal in cap site preference could be attributed to the relatively low binding affinity of *S.pombe* TBP to the TTAAGA motif, rather than to its rate of dissociation from the complex. Further, by constructing appropriate promoter mutants, we demonstrate that there is a good correlation between the nucleotide sequence of the TATA element (regardless of its location in the promoter) and the level of transcription. However, in human cells, it is the context in which the TATA element is placed, rather than its precise nucleotide sequence that appears to influence promoter activity. The data point to the possible existence of additional factor(s) in human cells that may serve to target TBP to nonconsensus TATA elements. Presumably, these factors are absent in *S.pombe*. This could be of significance in the investigation of mammalian transcription using the *S.pombe* system.

MATERIALS AND METHODS

Yeast strains and plasmids

S.pombe strain Ade 704 *h*⁺ (*leu1-32*, *his3-237*, *ura4-d18*) (43) was used throughout this investigation. The plasmid which contains a wild type (WT) Ad E2-early promoter fused to the reporter gene chloramphenicol acetyl transferase (CAT) has been described earlier (16,17). The WT promoter fused to CAT coding sequences was cloned into an autonomously replicating yeast shuttle vector, pURA4 (41). A 1.95 kb *Xho* I-*Xba* I fragment which contains the WT E2-early promoter, CAT coding sequences and the SV40 small-t splice and polyA signals (17) was first cloned into pUC19. By using the *Bam* HI and *Sph* I sites of the polylinker in this intermediate plasmid, the E2-early CAT gene was then cloned between *Bam* HI and *Sph* I sites of pURA4 (pU4E2CAT). In mutant E2-REV, the two TATA motifs were interchanged so that the +1 and -26 promoters contain TAAATT and TTAAGA sequences, respectively. In E2-CON, the TATA motif of the +1 promoter was replaced by a consensus TATA motif (Ad E1B TATA sequence), leaving the TATA motif between -59 and -52 unaltered. These mutations were first introduced into the E2-CAT WT plasmid by cloning oligonucleotides carrying the appropriate TATA motifs between the *Bss*H II (at -17) and *Hind* III sites (at -96). These were then transferred into pU4E2CAT background by replacing the *Bgl* II to *Sal* I fragment. The mutations in these constructs were confirmed by DNA sequence analysis.

RNA analysis

S.pombe cells were transformed with pU4E2CAT and the mutant derivatives according to the method of Beach and Nurse (43). Transformed colonies containing stably replicating plasmid were grown in minimal media to a density of 5×10^6 /ml. RNA was prepared from the transformed cells in mid-exponential phase as described (35). Total RNA from human cells was prepared using the guanidinium isothiocyanate protocol (44). Poly A⁺ RNA was used in the initial experiments whereas in the later experiments total RNA was used for primer extension analyses. Briefly, a 27 nucleotide (nt) long CAT primer corresponding to nts +110 to +136 (with reference to +1 cap site) was 5' end-labeled with T4 polynucleotide kinase using gamma ³²P ATP

(6000 Ci/mmol), gel purified and 10^5 cpm of this primer (equivalent to about 0.4 ng) was annealed to RNA (10 μ g poly A⁺ RNA; 50–100 μ g total RNA) in 30 μ l of hybridization buffer which contained 80% formamide, 40mM PIPES, pH 6.4, 400mM NaCl and 1mM EDTA, overnight at 30°C. The nucleic acids were precipitated and resuspended in a buffer containing 50mM Tris-HCl, pH 8.0, 50mM KCl, 5.0mM MgCl₂, 5.0mM DTT, 0.5mM dNTPs, 100 μ g/ml BSA, and 2.5 μ l/ml RNasin. Forty units of AMV reverse transcriptase were then added and the sample was incubated for 3 h at 42°C. The reaction was stopped by adding RNase A (40 μ g/ml) and EDTA (20mM) and incubating at 37°C for 30 min. The nucleic acids were phenol extracted, ethanol precipitated and resolved on a 6% DNA sequencing gel. *S.pombe* alcohol dehydrogenase mRNA (45) was quantitated under identical conditions using an appropriate primer to normalize the RNA levels. Similarly, the human RNAs were normalized by quantitating beta actin mRNA using a primer that was complementary to sequences extending from +78 to +105 of the human beta actin gene.

Bacterial expression and partial purification of TBP

pET vectors (46,47) encoding human and *S.pombe* TBP were expressed in *BL21 (DE3)pLysS* host bacteria and the TBPs partially purified using protocols published earlier (20,28) with some modifications. Expression was induced in bacterial cultures growing in logarithmic phase by the addition of IPTG to a final concentration of 0.4mM. Two cultures (2 liters each), one expressing human TBP and the other expressing *S.pombe* TBP were processed in parallel. Cells were harvested 3 h after induction at 37°C and suspended in 10ml buffer I (20mM Hepes, pH 8.0, containing 0.5M KCl, 20mM beta-mercaptoethanol, 1mM PMSF, 20 μ g/ml each of leupeptin, antipain and aprotinin and 10% glycerol) and lysed by sonication. The lysates were centrifuged (45K, 4°C, 50Ti rotor) to remove insoluble material and diluted with 4 volumes of buffer II (20mM Hepes, pH 8.0, containing 1 mM each of EDTA, PMSF and DTT, and 10% glycerol) and applied onto a 20ml DEAE-cellulose column equilibrated with 0.1M KCl in buffer II. The flowthroughs from the DEAE-cellulose columns were loaded onto 5ml heparin-agarose columns previously equilibrated with 0.1M KCl in buffer II. The columns were washed and then eluted successively with 0.25, 0.5 and 0.8 M KCl in buffer II. Eluted fractions were assayed for TBP activity in gel-shift assays using 5' end-labeled double-stranded oligonucleotides containing the adenovirus E1B TATA box sequence (28). Peak fractions eluted with 0.8M KCl were pooled together, dialyzed against 0.1M KCl in buffer II, clarified by centrifugation and used in all the experiments described in this paper.

Gel-shift assays

Gel-shift assays were carried out using a modification of the protocol described by Kao *et al* (28). Briefly, binding reactions were initiated by the addition of suitably diluted TBP preparation to approximately 0.2ng of 5' end-labeled probe (20–25,000 cpm) in a 20 μ l reaction containing 12mM Hepes, pH 8.0, 50mM KCl, 5mM MgCl₂, 1mM DTT, 50 μ g/ml BSA, 16 μ g/ml poly dI.dC and 10% glycerol. Reactions were incubated at 28°C for 25 min and run on a 5% native polyacrylamide gel in 1 $\times$ TBE buffer at 200V for 75 min at 4°C. After the run, gels were transferred onto a Whatman paper, dried and exposed for autoradiography.

Figure 1. Comparison of E2-early promoter activity in *S.pombe* versus HeLa cells by primer extension analysis: Poly A⁺ RNA was prepared from *S.pombe* cells stably transformed with pU4E2CAT plasmid and HeLa cells infected with an Ad variant harboring the E2-CAT construct. The RNA was annealed to a CAT primer complementary to sequences from +110 to +136 with respect to the +1 start site, reverse transcribed and analysed on a DNA sequencing gel. (A) The nucleotide sequence of the Ad E2-early promoter. The arrows denote the +1 and -26 cap sites of the dual promoter. The underlined sequences indicate the binding sites for TBP, E2F and ATF. (B) mRNAs transcribed from the two promoters in *S.pombe*. (C) mRNAs transcribed from the two promoters in HeLa cells. See reference 16 for further details of transcription of these promoters in HeLa cells. The products generated by primer extension are shown by arrowheads. The lane Markers contained an end-labeled *Hae* III digest of pBR322 DNA.

Off-rate analysis

Off-rate experiments were done based on a previously published procedure (48). Binding reactions were set up as described above, except that each reaction was scaled up to 140 μ l. After incubation at 28°C for 25 min, a 250-fold molar excess of unlabeled probe was added into the reaction and allowed to incubate for a further 60 min. During this 60 min period, 20 μ l aliquots were withdrawn at 10 min intervals, starting from 0 min (immediately after the addition of unlabeled probe) and loaded onto a running gel. The run was stopped 75 min after the last time point. To take into account any dissociation of the TBP-DNA complex during electrophoresis (10), control reactions were run in parallel. These reactions were set up exactly as described above except that no unlabeled probe was added. Off-rate was quantitated by measuring the radioactivity of the complexes excised from the dry gel followed by Cerenkov counting.

RESULTS

Adenovirus E2-early promoter is active in *S.pombe*

Previous studies by us and others have demonstrated that in human cells the Ad E2-early promoter is transcribed from the +1 start site about 20 fold more efficiently than from the -26 start site (15–18). We demonstrate here that this ratio is reversed in *S.pombe* cells (see below). To avoid the confusion that could arise by describing these promoters as major and minor as done earlier (15–18), the two promoters are now referred to as the +1 and -26 promoters. The nucleotide sequence of the Ad E2-early promoter with its dual start sites (+1 and -26) and the nonconsensus TATA sequence elements, TTAAGA and TAAATT, that control initiation at the two cap sites are shown in Figure 1A. For the sake of clarity and in conformity with the terminology adopted for the dual overlapping promoters above, the TTAAGA element (located between nts -29 and -22) which

controls the +1 cap site is referred to as the TBP (+1) motif; and the TAAATT element (located between nts -59 and -52) controlling the -26 cap site as the TBP (-26) motif. The other *cis* acting elements of the +1 promoter have been extensively characterized by us and others (see Figure 1A). These include two E2F sites, TTTCGCGC, present in an inverted orientation with respect to each other between nts -35 and -68 and an ATF site, TGACGTAG, between -68 and -77 (16,49-51). In mammalian cells this promoter is transactivated by the viral gene products E1A (reviewed in 52), and E4 (53), cellular gene product NFIL-6 (54) and several heterologous viral transactivators (55,56). Sequences of about 80 bp upstream from the +1 start site are both necessary and sufficient for the basal as well the induced transcription of the +1 promoter (16,17). However, not much is known about other *cis* elements that may influence -26 promoter activity. It must be pointed out here that the TBP (+1) motif overlaps the -26 cap site.

To determine if the E2-early promoter is expressed in fission yeast, *S.pombe* cells were transformed with the yeast shuttle vector pURA4 carrying the E2-early promoter CAT construct. We could detect considerable amounts of E2-early promoter driven CAT activity in the *S.pombe* transformants (data not shown). Having ascertained that the E2-early promoter is expressed in *S.pombe*, we proceeded to determine the individual contributions of each of the two overlapping promoters by quantitating transcription originating from the +1 and -26 promoters by primer extension analyses.

Relative efficiency with which the two promoters are transcribed is reversed in *S.pombe*

Poly A⁺ RNA was prepared from *S.pombe* cells transformed with pURA4 containing the E2-early CAT gene and annealed to a 5' end-labeled 27 nt-long primer complementary to nts +110 to +136 (with respect to +1 start site) of the CAT gene. The primer was then extended with reverse transcriptase and analysed on a DNA sequencing gel. This strategy can distinguish between the mRNAs transcribed from the two overlapping promoters (16). The mRNAs transcribed from the +1 promoter should generate a 136 nt-long product, whereas mRNAs transcribed from the -26 promoter should generate a product 162 nt-long. For comparison, we also determined the expression of the WT E2-early CAT construct cloned in the nonessential E3 region of a recombinant adenovirus lacking E1A in infected HeLa cells (16). As shown in Figure 1C, two cDNA products of the predicted sizes, 136 and 162 -nts long, which correspond to transcriptional initiation from the +1 and the -26 promoter, respectively, were identified in human cells. Densitometric scanning of the radioactive bands showed that the relative ratio of transcription from the +1 promoter to the -26 promoter is about 20:1. Reverse transcription of poly A⁺ RNA from *S.pombe* cells transformed with the E2-early CAT gene also produced two cDNA products of 136 and 162-nts similar to those seen in human cells (Figure 1B). Thus in *S.pombe*, the E2-early promoter initiated transcription from sites identical or very close to those detected in human cells (see Figure 2 for more accurate mapping of the start sites). Surprisingly, the relative ratio of transcription from the +1 promoter to the -26 promoter was reversed (compare WT lanes of Figure 1B and 1C). The +1 promoter whose TATA motif has a weak homology to the consensus TATA box was transcribed inefficiently in *S.pombe* whereas the -26 promoter whose TATA motif has a greater homology to the consensus TATA box was transcribed to much higher levels and the ratio

Figure 2. Role of nucleotide sequences of the TATA motifs in determining the levels of transcripts originating from the +1 and -26 cap sites of the E2-early promoter in *S.pombe*: (A) The nucleotide sequences of the E2-WT, E2-REV and E2-CON variants. For the sake of clarity, only the sequences in the vicinity of the TATA motifs and transcriptional start sites are shown. The three promoters differ only with respect to the boxed TATA motifs. Arrows above the WT promoter sequence indicate the start sites utilized in human cells. Arrows below indicate the start sites utilized in *S.pombe*, mapped in this study. (B) Primer extension analysis using total RNA prepared from *S.pombe* transformed with the promoter variants shown in A (lanes 1-3). Total RNA used for primer extension in lane 4 was prepared from human cells infected with an Ad variant expressing the WT E2-early-CAT construct. To determine the start sites precisely in these variants, a DNA sequence ladder (lanes G, A, T, C) representing the noncoding strand of the transcribed region was developed with the same CAT primer used for reverse transcription. The arrowheads indicate the position of RNA start sites of the E2-WT promoter in HeLa cells. Transcripts originating at the +1 and -26 regions (shown in brackets) were quantitated using Fujix Bioimage analyser. The CAT transcripts were normalized with respect to the levels of ADH mRNAs. (C) ADH transcripts in the total RNA samples (used in panel B) quantitated using an ADH primer complementary to sequences extending from +86 to +109 of the *S.pombe* ADH gene.

of the +1 to -26 promoter transcription was found to be about 1:20. Thus, these results show that *S.pombe* cells are capable of recognizing the two nonconsensus TATA motifs and initiating transcription roughly from the same cap sites that are used in human cells (15-18). However, the observation that *S.pombe* cells transcribe predominantly from the -26 promoter is most

striking. This led us to the obvious speculation that perhaps in *S.pombe*, the closer resemblance of the nucleotide sequence of the TATA motif of the -26 promoter to the consensus TATA box sequence could be responsible for this observation.

Transcriptional efficiency of the two overlapping E2-early promoters in *S.pombe* correlate with the nucleotide sequences in the TATA motif

Since the two cap sites of the E2-early promoter are controlled by two distinct, but overlapping transcriptional regulatory regions in human cells, it is very likely that the regulatory elements controlling the expression of this dual promoter in *S.pombe* overlap too. Furthermore, the TTAAGA sequence responsible for initiation of transcription at the $+1$ cap site overlaps with the nucleotide sequences of the -26 cap site (see Figure 1A). Thus, the unique architecture of this dual promoter precludes any extensive mutation of the TATA motif of the $+1$ promoter, for this would most certainly affect initiation controlled by the TATA element at the -26 cap site (Fig. 1A). With a view to retaining the overall promoter architecture, and to avoid possible complications that may arise due to the overlapping of the control elements mentioned above, we chose to mutate just the two TBP motifs of the promoter such that the TATA motif of the $+1$ promoter resembled the TATA motif of the -26 promoter and *vice versa*. Such an approach would permit us to address the role of the TATA motifs in E2-early promoter expression, specifically enabling us to ask if the precise sequence of nucleotides in the TATA motifs or the context in which the TATA motifs are placed in the promoter contribute to the efficiency of transcriptional initiation at a given cap site. To this end, we constructed the mutant E2-REV in which the TATA motifs of the $+1$ and -26 promoter were interchanged. This mutant contains the sequence TAAATT in the place of the sequence TTAAGA of the $+1$ promoter and the sequence TTAAGA in the place of the sequence TAAATT of the -26 promoter (Figure 2A). This construct with the swapped TATA motifs was transferred into pURA vector and *S.pombe* cells were transformed with the mutant clone. A large number of transformants were pooled, grown in liquid medium and total RNA was prepared as described in Materials and Methods. The relative efficiency of transcription from the $+1$ and -26 promoters was then assessed by a primer extension experiment as before, using total RNA. The radioactive bands from the transcripts generated in $+1$ and -26 region were quantitated using the Fujix Bioimage analyser and normalized to the ADH mRNA levels (Figure 2C; the origin of the band above the ADH mRNA was not investigated. It could represent a minor start site or it may be a nonspecific band). For the purpose of quantitation, all radioactive bands in the $+1$ and -26 region were included. To determine the start sites precisely in these variants, a DNA sequence ladder representing the noncoding strand of the transcribed region was developed using the same primer used in the primer extension experiments (lanes G, A, T, C in Fig. 2B). As expected, the relative ratio of transcription from the $+1$ and -26 cap sites in the WT promoter was about 1:20. (Note that the signals obtained in this experiment for the $+1$ and -26 transcripts are much weaker when compared to results obtained in Figure 1. This is because in the experiment shown in Figure 1, polyA⁺ RNA was used whereas in this experiment total RNA was used). In E2-REV in which the two TATA motifs were swapped, this ratio was also reversed and the $+1$ promoter became the major promoter. The -26 cap site of this mutant was utilized with an efficiency similar to that of

the $+1$ promoter of the WT. Clearly, the nucleotide sequence of the TATA motif is important in determining the efficiency of transcription in *S.pombe*. To further establish the importance of the nucleotide sequence, we constructed another mutant, E2-CON in which the TATA motif of $+1$ cap site of the WT promoter was replaced with the consensus motif, TATATA, derived from the Ad E1B promoter. This mutant retained the TATA motif of its -26 promoter unaltered (Figure 2A). When E2-CON was expressed in *S.pombe*, transcription from the $+1$ promoter was increased by about 75 fold when compared to that of E2-REV (compare lanes 2 and 3, Figure 2B) in which the TATA motif, TAAATT, controls the $+1$ promoter and by about 1400 fold when compared to that of E2-WT (compare lanes 1 and 3) in which the TATA motif, TTAAGA, controls the $+1$ promoter. In E2-CON, the RNA from the -26 cap site was also increased by about 3 fold when compared to that of the WT promoter. Note that the alterations in the E2-REV and E2-CON mutants are confined to nucleotide sequences from -29 to -22 and -59 to -52 . Such alterations are not expected to influence the interactions at other previously identified transcriptional control regions (in human cells). As all these mutants carry identical upstream activator elements (Figure 1A), the major reason that led to strikingly different levels of transcription from the $+1$ or -26 cap sites is the nucleotide sequence differences in the TATA elements.

It is noteworthy that in mutants E2-REV and E2-CON the RNA start sites are also changed. For example, in E2-REV, transcription from the $+1$ promoter starts at multiple initiation sites between $+2$ and $+6$. Similarly, in E2-CON, transcription from the $+1$ promoter starts at multiple initiation sites between $+3$ and $+9$. Likewise, in E2-CON, the -26 promoter initiated transcription at -25 and -27 positions. Two important features of these multiple initiation sites are that they are approximately about 30 bp downstream from the TATA motifs and all are purines except one. Both these features are typical of mammalian RNA polymerase II transcription (57).

Transcription from the $+1$ promoter in human cells does not appear to be critically dependent on the precise nucleotide sequence of the TATA element

Since the above studies demonstrated that the nucleotide sequence of the TATA box plays a critical role in determining the efficiency of transcription in *S.pombe*, we were interested to find out how the E2-REV and E2-CON variants of the Ad E2-early promoter would be expressed in human cells with respect to start site preference. Total RNA was isolated from human 293 cells 48 hours after transfection with the E2-early promoter constructs described in the preceding experiment and analyzed by primer extension as before (Figure 3). Transcripts originated predominantly from the $+1$ promoter, regardless of whether the nucleotide sequence of the TATA element located between nts -29 and -22 was TTAAGA (E2-WT), TAAATT (E2-REV) or TATATA (E2-CON). This was not unexpected because we do know from earlier reports that the $+1$ promoter with its atypical TTAAGA motif is the major promoter in human cells, despite the availability of a closer to consensus TAAATT motif in the -26 promoter (15–18). Our data corroborate the view that nucleotide sequence requirements for TATA boxes are relatively less stringent for the expression of the E2-early promoter in human cells. However, the observation that improving the quality of the TATA box of the $+1$ promoter did not significantly enhance transcription of this promoter is

Figure 3. Expression of the E2-early promoter constructs (described in Figure 2) in human 293 cells: 293 cells were transfected with E2-WT, E2-REV or E2-CON at 80% confluency using the calcium phosphate protocol (44). Total RNA was prepared 48 hours after transfection and analyzed by primer extension. In lanes 2, 3, 4 and 9 total RNA from transfected cells was analyzed using the CAT primer (described in Figure 1). The CAT transcripts originating from the +1 and -26 cap sites of E2-WT in infected 293 cells are shown in lane 1. RNA levels were normalized by quantitating beta actin RNA in the transfected cells (lanes 5-8) using a primer complementary to sequences extending from +78 to +105 of the human beta actin gene (16). The arrowheads indicate the positions of the reverse transcribed products.

intriguing (compare lanes 2,3 and 4). Bioimage analysis of the gel shown in Figure 3 revealed the levels of transcripts originating at the +1 cap sites of E2-WT, E2-REV and E2-CON constructs to be very similar to each other. Previous studies have shown that the E1A protein transactivates the E1A sensitive promoters including E2-early promoter by interacting with TBP (58-60). It is possible that the preferential utilization of +1 promoter in these three promoter mutants could be due to the E1A protein expressed in the 293 cells (61). To rule out the effect of E1A on the start site preference, the above experiment was repeated in HeLa cells which lack E1A. Although the signals were considerably weaker when compared to that of 293 cells, the ratio of the +1 promoter to -26 promoter utilization in HeLa cells was very similar to that of 293 cells (data not shown) and it was the +1 promoter that was preferentially transcribed in all three constructs (data not shown). These results indicate that the human RNA polymerase II machinery is able to utilize a wide range of TATA boxes equally efficiently. We once again note a slight change in the start site for the +1 promoter of the E2-CON mutant. In E2-REV and E2-CON the relative levels of transcription at the -26 cap site was reduced compared to that of E2-WT. (Compare lanes 3 and 4 with lane 2 in Figure 3).

Both human and *S.pombe* TBP bind to the TAAATT motif better than to the TTAAGA motif

One explanation for the relative difference in the utilization of the two promoters may be due to differences in the binding efficiency of *S.pombe* and human TBPs to the two TATA-like motifs. For example, the human TBP would be expected to bind to TTAAGA motif better than to the TAAATT motif whereas the converse would be true for *S.pombe* TBP. We therefore studied the binding efficiency of *S.pombe* and human TBPs to the two TATA-like motifs *in vitro* using gel-shift assays.

Figure 4. Comparison of the binding of *S.pombe* and human TBPs to the TTAAGA and TAAATT motifs of the Ad E2-early promoter: Gel-shift assays were carried out using a constant amount of protein at varying concentrations of radiolabeled TATA box containing duplex oligonucleotides as described by Baker *et al* (62). Numbers above the lanes represent the amount of probe taken (in cpm $\times$ 1000) in each reaction. The specific activities of the different probes used was typically in the range of $1.5-3.0 \times 10^5$ cpm/ng. C: control reaction in which protein was omitted. Lane numbers are indicated below the panels. (A) Binding of *S.pombe* TBP to different TATA boxes. (B) Same as A, but with human TBP. (C) Nucleotide sequences of the TATA box containing oligonucleotides used in this experiment (all the oligonucleotides were designed with *Sal* I overhangs at their 5' ends). The arrowheads indicate the TBP-DNA complexes. These were cut and counted directly to quantitate relative binding efficiency.

Recombinant TBPs were expressed in *E.coli*, and the proteins were partially purified by successive chromatography on DEAE-cellulose and heparin-agarose columns. The fractions were monitored in gel-shift assays for binding activity using the Ad E1B TATA motif as a probe. The active fractions from the heparin-agarose columns were pooled and used in gel-shift assays with duplex oligonucleotides containing different TBP motifs. The specificity of the interaction in each case was ascertained in competition experiments (data not shown).

To determine the relative affinity of *S.pombe* and human TBPs for the different TATA elements, we utilized the gel-shift assays as described by Baker *et al* (62) in which the DNA binding activity of a constant concentration of protein is analyzed at varying concentrations of radiolabeled probe DNA. Since the impurities in our partially purified TBP preparations and their effects on the binding are not known, the above approach which titrates a constant amount of TBP as a function of DNA ligand concentration should permit a more reliable and accurate assessment of the relative binding efficiencies. Gel-shift assays were carried out with the following 4 oligonucleotides (Figure 4C): (i), TATA/+1 which contains the TATA motif of

the +1 promoter (nucleotide sequence from -37 to -15); (ii), TATA/-26 which contains the TATA motif of the -26 promoter (nucleotide sequence from -66 to -44); (iii), TATA/-26R in which the TATA motif of the +1 promoter, TTAAGAGT, is substituted by the TATA motif of the -26 promoter TTAAATTT and (iv), TATA/CON in which the TATA motif of the +1 promoter is substituted by a consensus TATA motif (E1B promoter TATA box). The oligonucleotides TATA/+1, TATA/-26R and TATA/CON are identical with respect to the sequences flanking their TATA motifs (nucleotide sequences underlined in Figure 4C). Therefore when these three oligonucleotides are used as probes in the gel-shift assays, possible contribution by the flanking sequences to differences in binding can be ruled out.

Both *S.pombe* (Figure 4A) and human (Figure 4B) TBPs bound with greater efficiency to the TATA motif of the -26 promoter (TATA/-26) compared to that of the +1 promoter (TATA/+1). The results with *S.pombe* protein were striking. When the radiolabeled bands (shown by arrows) were excised and quantitated by Cerenkov counting, it was found that the TBP of *S.pombe* bound to TAAATT about 5–8 fold better than it did to TTAAGA (compare lanes 2–5 with 7–10, Figure 4A). The difference in the binding efficiency between the two TATA motifs was about 3 to 5 fold for the human TBP (compare lanes 2–5 with 7–10, Figure 4B). As these two oligonucleotides, TATA/+1 and TATA/-26 differ not only in their TATA box sequences but also in the sequences flanking their TATA boxes (Figure 4C), these data are not sufficient to attribute the observed differences in binding to the differences in the TATA boxes alone. However, a comparison of the binding affinities of these two TATA boxes when they are placed in identical backgrounds (compare TATA/+1 with TATA/-26R, Figure 4C) shows unambiguously that the TAAATT element does indeed display better binding to the TBPs. *S.pombe* TBP bound to the TAAATT element (TATA/-26R) with 7–12 fold greater efficiency than to the TTAAGA element (TATA/+1) (compare lanes 2–5 with 12–15, Figure 4A) with identical flanking sequences. Similarly, human TBP bound about 5–8 fold more efficiently to the TAAATT element (TATA/-26R) as compared to the TTAAGA element (TATA/+1) (compare lanes 2–5 with 12–15, Figure 4B). Thus, we conclude that it is the quality of the TATA element (*i.e.* the degree of homology to the consensus sequence) which determines the efficiency of binding to the TBPs. We do note that the TBP–DNA interaction is slightly better when the TAAATT motif is present in the TATA/-26R background rather than in the TATA/-26 background, reflecting that sequences flanking the TATA box might influence binding. However, this effect is not significant as the binding strengths of TATA/-26 and TATA/-26R vary within 2–3 folds of each other for both human as well as *S.pombe* TBPs. In terms of binding affinity, the TAAATT motif was quite comparable to the TATATA motif (compare TATA/-26 and TATA/-26R with TATA/CON). Thus, it is clear from these gel-shift experiments that the TAAATT sequence of the -26 promoter is better than the TTAAGA sequence of the +1 promoter and equivalent to the consensus sequence TATATA with respect to binding to TBP *in vitro*.

TBP dissociates from TTAAGA and TAAATT motifs at similar rates

Another possible reason for the large difference in the efficiency with which the +1 and -26 promoters are transcribed may be

Figure 5. Off-rate analysis of the human and *S.pombe* TBP complexes. TBP–DNA complexes were allowed to form under standard gel-shift assay conditions followed by the addition of a 250 fold molar excess of unlabeled homologous probe. Aliquots were then withdrawn at 10 minute intervals and loaded onto a running native polyacrylamide gel. The numbers above the lanes indicate the time in minutes (after the addition of unlabeled probe DNA) when the lanes were loaded. The specific TBP and the oligonucleotide used are indicated at the bottom of each panel. TTAAGA and TAAATT correspond to TATA/+1 and TATA/-26 of Figure 4C. The arrowheads indicate the TBP–DNA complexes. For quantitation, the complexes were cut out from the gel and counted directly. The values from panels A and B are plotted in panel C. Likewise, the values from panels D and E are plotted in panel F. Control experiments were run in parallel in which the addition of unlabeled oligonucleotide was omitted. This was done to quantitate any possible dissociation of the TBP–DNA complex during electrophoresis (10). There was no significant dissociation of the complexes in these control experiments (data not shown).

due to the differences in the dissociation rates of the complexes formed between the TBPs and the two TATA-like motifs. It could be argued that in *S.pombe*, for example, a relatively rapid dissociation of TBP from the TBP–TTAAGA complex compared to the TBP–TAAATT complex might be responsible for the preferred utilization of the -26 promoter. We therefore determined the off-rate for TTAAGA– and TAAATT–TBP complexes with human and *S.pombe* TBPs as described in Materials and Methods. Results presented in Figure 5 show that both human and *S.pombe* TBP complexes with the two different TATA-like motifs were stable. The half-lives of the complexes formed with TTAAGA and TAAATT with human protein were 30 and 24 minutes, respectively, whereas with *S.pombe* protein the half lives were 14 and 20 minutes, respectively. Control reactions were run in parallel to ascertain the degree of dissociation of the TBP–DNA complex as reported by Hoopes *et al* (10). No significant dissociation occurred during electrophoresis under the experimental conditions (data not shown). The stability of these complexes is lower than that observed earlier for *S.cerevisiae* TBP bound to the TATA motif of the Ad MLP (63). This difference is most likely attributable to the nature of the two protein preparations and therefore direct comparison is difficult. Nonetheless, the complexes shown in

Figure 5 appear to be considerably stable. These results show that both TAAATT and TTAAGA motifs dissociated from the complex at similar rates indicating that difference in the efficiency of transcription between the two promoters may not be due to a difference in the rate of dissociation of the complexes.

DISCUSSION

In this study we have demonstrated, for the first time, that the human Ad E2-early promoter is active in *S.pombe* and that both the cap sites of this dual promoter are functional. However, the surprising observation is that in contrast to mammalian cells, this efficient transcription is initiated not from the +1 promoter which is a very weak promoter in *S.pombe* but rather from the -26 promoter. As in human cells (15-18), transcription of these promoters in *S.pombe* is also dependent on the two atypical TATA motifs because efficiency of transcription from these two cap sites correlated well with the nucleotide sequences of the TATA motifs (see below).

The sequence TAAATT located between -59 and -52 that controls the -26 cap site resembles more closely the consensus TATA motif compared to the TTAAGA sequence located between -29 and -22 which serves as a TATA element for the +1 promoter. Analysis of the mutant in which the TATA motifs are interchanged with each other (E2-REV) and the mutant in which the TTAAGA motif is substituted by the E1B TATA motif (E2-CON) provide direct evidence that the precise nucleotide sequence of the TATA element and not the context in which these elements are located is the major determinant factor responsible for the difference in the levels of transcription of these promoters in *S.pombe*. When the TTAAGA motif is replaced by the TAAATT or TATATA motifs, transcription of the +1 promoter is increased by 20 and 1400 fold, respectively. Similarly, when TAAATT motif of the -26 promoter is replaced with TTAAGA motif, transcription decreases by about 20 fold. A recent observation that the TATA motif of the +1 promoter did not function efficiently in *S.pombe* is also consistent with our findings (37).

Recently it has been shown that sequences flanking the cap site as well as sequences further downstream of TATA box lacking promoters appear to be important for mediating the stable association of TBP with the promoter (64,65). These cap site flanking sequences termed as initiator (Inr) elements which have been identified in many TATA box lacking polymerase II promoters determine the precise site of initiation (64). Currently we do not know whether sequences flanking the +1 and -26 cap sites of the E2-early promoter have any role as Inr elements. Alterations in the +1 cap site sequences do not have any major effect on transcriptional initiation at this cap site in mammalian cells (17,18). The unique architecture of the Ad E2-early promoter presents a situation wherein the TTAAGA element (controlling the initiation of transcription from the +1 cap site) overlaps the -26 cap site. Therefore, the role of Inr elements, if any, on the efficiency of transcription at the -26 cap site is difficult to evaluate.

We examined whether the preferential transcription from the -26 promoter in *S.pombe* and from the +1 promoter in human cells could be due to a difference in the affinity of these TATA-type motifs to bind to TBP from these two organisms or due to their dissociation rates or both. That is, the human TBP would bind to TTAAGA motif with higher affinity compared to TAAATT motif whereas *S.pombe* TBP would bind to TAAATT

motif more efficiently compared to the TTAAGA motif. The TBP binding studies show that both human and *S.pombe* TBP bound to the TTAAGA motif less efficiently compared to the TAAATT motif; this difference is particularly striking in the case of *S.pombe*. The TBP of *S.pombe* bound to TAAATT motif 7-12 fold more efficiently when compared to the TTAAGA motif, whereas human TBP bound to TAAATT sequence about 5-8 fold more efficiently when compared to the TTAAGA sequence. In both cases the dissociation rate of the TBP from the TBP-TATA motif complexes was found to be very similar indicating that once TBP binds to these motifs, the resulting complexes are quite stable. Thus, it seems likely that the affinity of TBP for the two TATA motifs might be one important factor in determining the efficiency of transcription in *S.pombe*. We, however, note the discrepancy that *S.pombe* TBP binds to a consensus TATA sequence as well as to the TAAATT motif. Yet, transcription of the +1 promoter increases dramatically when the consensus TATA motif replaces the TAAATT sequence. It is possible that the binding differences may be too small to be detected in gel-shift assays and yet may have a major impact on protein-protein interactions between upstream activator proteins and general transcription factors in forming a stable initiation complex *in vivo*. Both TATA-like motifs of the E2-early promoter deviate from the consensus TATA sequence. It is conceivable that the TAAATT sequence is tolerated to a higher degree than the TTAAGA sequence by the *S.pombe* basal transcription machinery. Thus, a major factor that may contribute to the preferential transcription of the -26 promoter in *S.pombe* may be the higher binding affinity of its TBP to the TAAATT motif. Thus, a good TATA box is crucial for the efficient expression of a mammalian (or *S.pombe*) promoter in fission yeast. This observation may be significant for studies on the regulation of mammalian promoters in *S.pombe*.

In mammalian cells the +1 promoter is transcribed about 20 fold more efficiently despite the fact that the sequence motif TTAAGA that serves as the TATA box for this promoter binds to TBP less efficiently than does the -26 promoter specific TATA box in gel-shift assays (Figure 4). Further, replacing the atypical TTAAGA element with either TAAATT or TATATA did not significantly enhance transcription from the +1 promoter in human cells (Figure 3), emphasizing the ability of the mammalian transcription machinery to efficiently utilize nonconsensus TATA elements. Perhaps, the context in which a functional TATA element (despite a nonconsensus sequence) is located in the promoter is of greater importance in determining the efficiency of transcription in human cells. This is in striking contrast to the situation in *S.pombe* where the precise nucleotide sequence of the TATA element has a profound effect on the levels of transcription (Figure 2). It is apparent that subtle differences exist between the basal transcription machineries of the two systems. It is quite likely that the upstream regulatory factors are functioning differently in the two systems. Though the mammalian upstream factors that regulate the E2-early promoter are better understood (16,49-51), currently not much is known regarding their counterparts in *S.pombe*. There is a homolog of mammalian ATF in *S.pombe* (66). We have preliminary evidence that there is an E2F-related activity in *S.pombe* which is quite distinct from mammalian E2F (67). Additionally, the human and *S.pombe* TBPs display distinct differences. While the native yeast TBP is a monomeric protein, human TBP is an integral component of multisubunit complex (11,12). It is speculated that yeast TBP may be associated with additional factors analogous

to the human TAFs. However, there is so far no evidence that *S.pombe* TBP may function as a multisubunit complex.

Why do these two TATA motifs behave differently in two different systems? While a definite answer is yet to be found, we can speculate on some possibilities. For example, recently it has been shown that TBP interacts with a class of intermediary proteins called coactivators or adaptors in the assembly of the initiation complex (11,12). These coactivators are believed to bridge the TBP with upstream gene specific activators. It is conceivable that coactivators that bridge the upstream gene specific activators of the +1 promoter and the TBP bound to its cognate TATA motif may be more effective than the coactivators that participate in the activation of the -26 promoter. Another possibility that can be envisaged is that the TBP when bound to two different TATA-like elements may undergo different conformational changes. Recently, Berk and coworkers showed that the amino terminal region of *S.cerevisiae* TBP undergoes a conformational change upon DNA binding (68). As the amino terminal domains of *S.pombe* TBP and human TBP are very different with respect to size, sequence and composition, it is possible that the two TATA-like sequences of the E2-early promoter may induce two different conformational changes in TBP which may consequently be able to interact with upstream activators or coactivators with different efficiencies. These subtle differences may lead to widely different levels of transcription from the two promoters. Such a mechanism is attractive and consistent with the evidence that the TATA motifs in eucaryotic cells are functionally heterogeneous (16,58,68-70). The TBP contains a highly conserved 180 amino acid C-terminal DNA binding domain and a nonconserved amino terminal region of variable length (19-30). Genetic and biochemical analysis of the TBP indicate its function to be species specific (29,71,72). However, in addition to the divergent N-terminus, differences scattered throughout the conserved core domain account for the species variations in the function of this transcription factor (71,72).

In conclusion, we have analyzed the role of TATA elements in mammalian promoter expression in *S.pombe* and demonstrate that there are subtle, yet unique differences between *S.pombe* and human transcriptional machineries. While the nucleotide sequence of the TATA element is crucial for expression in *S.pombe*, the context in which the TATA element is placed appears to be critical for promoter activity in human cells. Apparently, *S.pombe* transcription machinery is a relatively simpler system relying rather heavily on the inherent TATA sequence to initiate transcription (see Figure 2). By comparison, the mammalian transcription machinery is quite evolved and consequently more complex. The data presented in this study are consistent with the idea that mammalian cells may possess one or more factors that facilitate efficient transcription of promoters with nonconsensus TATA motifs, whereas *S.pombe* may lack these factors.

ACKNOWLEDGEMENTS

This work was supported by Public Health Service grants AI 20156 and AI18029 from the National Institutes of Health. We thank N.Jones for introducing us to the *S.pombe* system, M.Furtado and Urmila Shinde for help in certain experiments, P.Liebermann, A.J.Berk, A.Hoffman and R.G.Roeder for the TBP expression plasmids, Jo Ann Wise for providing *S.pombe* growth protocols and John Benson for critical reading of the manuscript.

REFERENCES

- Guarente, L. (1988) *Cell*, **52**, 303-305.
- Struhl, K. (1989) *Ann. Rev. Biochem.*, **58**, 1051-1077.
- Greenblatt, J. (1991) *Cell*, **66**, 1067-1070.
- Sawadogo, M., and Sentenac, A. (1990) *Ann. Rev. Biochem.*, **59**, 711-754.
- Struhl, K. (1987), *Cell*, **49**, 295-297.
- Nakajima, N., Horikoshi, M. and Roeder, R. G. (1988), *Mol. Cell. Biol.*, **8**, 4028-4040.
- Horikoshi, M., Yamamoto, T., Ohkuma, Y., Weil, P. A. and Roeder, R. G. (1990), *Cell*, **61**, 1171-1178.
- Starr, D. B. and Hawley, D. K. (1991), *Cell*, **67**, 1231-1240.
- Lee, D. K. Horikoshi, M. and Roeder, R. G. (1991), *Cell*, **67**, 1241-1250.
- Hoopes, B. C. LeBlanc, J. F. and Hawley, D. K. (1992), *J. Biol. Chem.*, **267**, 11539-11547.
- Pugh, B. F. and Tjian, R. (1991), *Genes and Dev.*, **5**, 1935-1945.
- Tanese, N., Pugh, B. F. and Tjian, R. (1991), *Genes and Dev.*, **5**, 2212-2224.
- Zhou, Q., Lieberman, P. M., Boyer, T. G. and Berk, A. J. (1992), *Genes and Dev.*, **6**, 1964-1974.
- Sharp, P. A. (1992), *Cell*, **68**, 819-821.
- Mathis, D. J., Elkaim, R., Keding, C., Sassone-Corsi, P. and Chambon, P. (1981) *Proc. Natl. Acad. Sci. U.S.A.*, **78**, 7383-7387.
- Manohar, C. F., Kratochvil, J. and Thimmapaya, B. (1990), *J. Virol.*, **64**, 2457-2466.
- Murthy, S. C. S., Bhat, G. P., and Thimmapaya, B. (1985), *Proc. Natl. Acad. Sci. U. S. A.*, **82**, 2230-2234.
- Zajchowski, D. A., Boeuf, H. and Keding, C. (1985), *EMBO J.*, **4**, 1293-1300.
- Cavallini, B., Faus, I., Matthes, H., Chipoulet, J. M. Winsor, B. Egly, J. M. and Chambon, P. (1989) *Proc. Natl. Acad. Sci. U.S.A.*, **86**, 9803-9807.
- Hoffmann, A., Horikoshi, M. Wang, C. K., Schroeder, S., Weil, P. A. and Roeder, R. G. (1990), *Genes and Dev.*, **4**, 1141-1148.
- Fikes, J. D., Becker, D. M., Winston, F. and Guarente, L. (1990), *Nature (London)*, **346**, 291-294.
- Eisenmann, D. M., Dollard, C. and Winston, F. (1989), *Cell*, **58**, 1183-1191.
- Hahn, S., Buratowski, S. Sharp, P. A. and Guarente, L. (1989), *Cell*, **58**, 1173-1181.
- Schmidt, M. C., Kao, C. C., Pei, R. and Berk, A. J. (1989), *Proc. Natl. Acad. Sci. U.S.A.*, **86**, 7785-7789.
- Hoey, T., Dynlacht, B. D., Peterson, M. G., Pugh, B. F. and Tjian, R. (1990), *Cell*, **61**, 1179-1186.
- Gasch, A., Hoffmann, A., Horikoshi, M., Roeder, R. G. and Chua, N. H. (1990), *Nature (London)*, **346**, 390-394.
- Hoffmann, A., Sinn, E., Yamamoto, T., Wang, J., Roy, A. Horikoshi, M. and Roeder, R. G. (1990), *Nature (London)*, **346**, 387-390.
- Kao, C. C., Lieberman, P. M., Schmidt, M. C., Zhou, Q., Pei, R. and Berk, A. J. (1990), *Science*, **248**, 1646-1650.
- Peterson, M. G., Tanese, N., Pugh, B. F. and Tjian, R. (1990), *Science*, **248**, 1625-1630.
- Haas, M. M. and Feix, G. (1992), *FEBS Lett.*, **301**, 294-298.
- Strubin, M. and Struhl, K. (1992), *Cell*, **68**, 721-730.
- Hahn, S., Buratowski, S., Sharp, P. A. and Guarente, L. (1989), *Proc. Natl. Acad. Sci. U.S.A.*, **86**, 5718-5722.
- Harbury, P. A. B. and Struhl, K. (1989), *Mol. Cell. Biol.*, **9**, 5298-5304.
- Russell, P. and Nurse, P. (1986), *Cell*, **45**, 781-782.
- Kaufman, N. F., Simanis, V. and Nurse, P. (1985), *Nature (London)*, **318**, 78-80.
- Lee, M. G. and Nurse, P. (1987), *Nature (London)*, **327**, 31-35.
- Prentice, H. L. and Kingston, R. E. (1992), *Nucl. Acids Res.*, **20**, 3383-3390.
- Russell, P. (1985), *Gene*, **40**, 125-130.
- Toyama, R. and Okayama, H. (1990), *FEBS Lett.*, **268**, 217-221.
- Matsumoto, S. and Yanagida, M. (1985), *EMBO J.*, **4**, 3531-3538.
- Jones, R. H., Moreno, S., Nurse, P. and Jones, N. C. (1988) *Cell*, **53**, 659-667.
- Losson, R., Fuchs, R. P. R. and Lacroute, F. (1985), *J. Mol. Biol.*, **185**, 65-81.
- Beach, D. and Nurse, P. (1981) *Nature (London)*, **290**, 140-142.
- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Smith, J. A., Seidman, J. G. and Struhl, K. (1987), *Current protocols in molecular biology*, Greene Publishing Associates and John Wiley and Sons, New York.
- Russell, P. R. and Hall, B. D. (1983), *J. Biol. Chem.*, **258**, 143-149.

46. Rosenberg, A. H., Lade, B. N., Chui, D. S., Lin, S.-W., Dunn, J. J. and Studier, F. W. (1987), *Gene*, **56**, 125–135.
47. Studier, F. W. and Moffatt, B. A. (1986), *J. Mol. Biol.*, **189**, 113–130.
48. Chodosh, L. A., Carthew, R. W. and Sharp, P. A. (1986), *Mol. Cell. Biol.*, **6**, 4723–4733.
49. SivaRaman, L., Subramanian, S. Thimmapaya, B. (1986), *Proc. Natl. Acad. Sci. U. S. A.*, **83**, 5914–5918.
50. SivaRaman, L. and Thimmapaya, B. (1987), *Proc. Natl. Acad. Sci. U. S. A.*, **84**, 6112–6116.
51. Yee, A. S., Raychaudhuri, R., Jakoi, L. and Nevins, J. R. (1988), *Mol. Cell. Biol.*, **9**, 578–585.
52. Flint, J. and Shenk, T. (1989), *Annu. Rev. Genet.*, **23**, 141–161.
53. Hardy, S., Engel, D. A. and Shenk, T. (1989), *Genes and Dev.*, **3**, 1062–1074.
54. Spengel, J. M., Hsu, W., Akira, S., Thimmapaya, B., Kishimoto, T. and Chen-Kiang, S. (1992), *J. Virol.*, **66**, 1021–1030.
55. Loeken, M. R. and Brady, J. (1989), *J. Biol. Chem.*, **264**, 6572–6579.
56. Phelps, W. C., Yee, C. L., Munger, K. and Howley, P. M. (1988), *Cell*, **53**, 539–547.
57. Breathnach, R. and Chambon, P. (1981), *Ann. Rev. Biochem.*, **50**, 349–383.
58. Wu, L., Rosser, D. S. E., Schmidt, M. C. and Berk, A. J. (1987), *Nature (London)*, **326**, 512–515.
59. Lee, W. S., Kao, C. C., Bryant, G. O., Liu, X. and Berk, A. J. (1991), *Cell*, **67**, 365–376.
60. Horikoshi, N., Maguire, K., Kralli, A., Maldonado, E., Reinberg, D. Weinman, R. (1991), *Proc. Natl. Acad. Sci. U. S. A.*, **88**, 5124–5128.
61. Graham, F. L., Smiley, J., Russell, U. C. and Nairu, R. (1977), *J. Gen. Virol.*, **36**, 59–72.
62. Baker, R. E., Gabrielsen, O. and Hall, B. D. (1986), *J. Biol. Chem.*, **261**, 5275–5282.
63. Schmidt, M. C., Zhou, Q. and Berk, A. J. (1989), *Mol. Cell. Biol.*, **9**, 3299–3307.
64. Smale, S. T. and Baltimore, D. (1989), *Cell*, **57**, 103–113.
65. Purnell, B. A. and Gilmour, D. S. (1993), *Mol. Cell. Biol.*, **13**, 2593–2603.
66. Jones, R. H. and Jones, N. C. (1989), *Proc. Natl. Acad. Sci. U. S. A.*, **86**, 2176–2180.
67. Malhotra, P., Manohar, C. F., Swaminathan, S., Toyama, R., Dhar, R., Reichel, R. and Thimmapaya, B. (Submitted).
68. Lieberman, P. M., Schmidt, M. C. Kao, C. C. and Berk, A. J. (1991), *Mol. Cell. Biol.*, **11**, 63–74.
69. Simon, M. C., Fisch, T. M., Benecke, B. J., Nevins, J. R. and Heintz, N. (1988), *Cell*, **52**, 723–729.
70. Leong, K., Brunet, L. and Berk, A. J. (1988), *Mol. Cell. Biol.*, **8**, 1765–1774.
71. Cormack, B. P., Strubin, M. Ponticelli, A. S. and Struhl, K. (1991), *Cell*, **65**, 341–348.
72. Gill, G., and Tjian, R. (1991), *Cell*, **65**, 333–340.