A novel TATA-box-binding factor from the silk glands of the mulberry silkworm, *Bombyx mori*

Lakshmi SRINIVASAN and Karumathil P. GOPINATHAN

Microbiology and Cell Biology Department, Indian Institute of Science, Bangalore-560012, India

The presence of one or more TATATAA motifs in the flanking sequences of individual members of a multi-gene tRNA\(^{\text{I/G}}\) family from the mulberry silkworm, *Bombyx mori*, negatively modulated the transcription of the gene copies. Characterization of proteins from posterior silk gland nuclear extracts, binding to the TATA TAA motif, identified a novel 43 kD protein, designated here as P43 TATA-box-binding factor (TBF). The protein was purified to homogeneity. P43 TBF binding was highly sequence-specific and showed a 100-fold-higher affinity for binding than the TATA-box-binding protein (TBP). The protein also showed binding to the TATAAA sequence of the actin\(^{\text{5C}}\) promoter. P43 TBF inhibited transcription of all the tRNA genes examined, as well as RNA polymerase II transcription from the actin\(^{\text{5C}}\) promoter. The amino acid sequence of eleven peptides generated from P43 TBF did not share homology with proteins that bind the TATA box, such as TBP, TRF (TBP-related factor) or TLFs (TBP-like factors) reported from other sources. Inhibition of transcription of tRNA genes by P43 TBF could not be reversed by TBP. The inhibitory effect appeared to be exerted through sequestration of the associated transcription factors.

Key words: RNA polymerase III, TATA binding protein, tRNA transcription, transcriptional regulation.

INTRODUCTION

The promoters for eukaryotic tRNA genes consist of two essential and highly conserved regions of approx. 10 bp each, the A- and B-boxes, located within the coding region and generally separated by 30–40 bp [1]. The prevailing models for tRNA gene transcription invoke the function of two multi-subunit transcription factors, TFIIIB and TFIIIC. The binding of TFIIIC to the A and B boxes leads to the binding of TFIIH immediately upstream of the tRNA-coding region, which in turn recruits RNA polymerase III (pol III) to initiate transcription at the start site [2–4]. The TATA-binding protein (TBP) is recruited to the transcription initiation complex as a component of TFIIH. Whereas the A- and B-blocks constitute the essential promoter elements of a tRNA gene, variations in the 5′ flanking sequences exert modulatory effects on their transcription in yeast [5–7], fruit flies [8,9], silkworms [10–18], frogs [19–21] and humans [22,23]. There is little or no sequence homology conserved between the 5′ flanking sequences of different tRNA genes, including even those gene copies encoding the same tRNA ‘isoacceptor’ species. However, the presence of A/T-rich sequences within 20–30 bp of the transcription start site of several tRNA genes occurs consistently. Comparison of the upstream regions of 12 individual tRNAs from yeast revealed an average of 68 %, A/T content between nt −1 and −40 [24]. The upstream regions of three copies of human tRNA\(^{\text{Ure}}\) contain 70–90 %, A/T content [22]. The members of tRNA\(^{\text{I/G}}\) family from *Bombyx mori* possess 63–68 %, A/T content within the 100 bp 5′ upstream sequences [14]. Transcription of tRNA\(^{\text{I/G}}\) and tRNA\(^{\text{I/G}}\) from *B. mori* are positively regulated by the A/T-rich motifs present in the immediate upstream to −40 nt of the coding region [14,25,26].

The tRNA population of differentiated cells is adapted to meet their special protein-synthetic demands, almost exclusively through the transcriptional regulation of the tRNA genes. A classic example of this functional adaptation of tRNA population occurs in the silk glands of the mulberry silkworm *B. mori*. Fibroin, the major constituent protein of the silk fibre, is synthesized at very high levels in the posterior silk glands during the fifth larval instar. The unusual amino acid composition of fibroin H chain (46 %, glycine, 33 %, alanine, 12 %, serine and 5 %, tyrosine) is reflected in the distribution of tRNAs in the posterior silk gland, the site of protein synthesis, to achieve optimized fibroin production [27,28]. The large 15 kb mRNA encoding fibroin H comprises 2415 codons for glycine [29], and of these 1352 correspond to GGC. These are base-paired with tRNA\(^{\text{Gly}}\). In *B. mori*, tRNA\(^{\text{Gly}}\) constitutes a multi-gene family with an estimated 20 copies, and 11 of these have been characterized previously by our group [13,14,16,17]. On the basis of transcription levels in *in vitro* posterior silk gland nuclear extracts, these copies of tRNA\(^{\text{Gly}}\) could be grouped into three distinct classes. Some of them were highly transcribed (tRNA\(^{\text{Gly}}\) 1 and 11), some were very poorly transcribed (tRNA\(^{\text{Gly}}\) 6, 7, 8, 9 and 10) and the rest were transcribed to intermediaries levels (tRNA\(^{\text{Gly}}\) 2, 3, 4 and 5). The *in vitro* transcription of these genes in *B. mori* derived BmN cell lines also followed essentially the same pattern [30,31]. Since all these gene copies have identical coding regions (and consequently harboured the same identical coding regions), the obvious candidates for mediating the differential transcription are the 3′ and 5′ flanking regions. In fact, sequences present far upstream of the coding region (up to 1 kb away) also modulated transcription of these tRNA\(^{\text{Gly}}\) copies [13,16,17].

A negative regulatory element present between −220 to −300 nt upstream of the tRNA coding sequence in tRNA\(^{\text{Gly}}\) has been identified to be a TATATAA motif located at −276 nt in the highly transcribed copy of tRNA\(^{\text{Gly}}\)-1 [13]. Multiple copies of this

Abbreviations used: TBF, TATA-box-binding factor; TBP, TATA-box-binding protein; TRF, TBP-related factor; TLF, TBP-like factor; TAF, TBP-associated factor; pol (II/III), RNA polymerase II and III respectively; DTT, dithiothreitol; EMSA, electrophoretic mobility shift assay; RK fragment, EcoRI-KpnI fragment.

1 To whom correspondence should be addressed (e-mail kpg@mcbl.iisc.ernet.in).
sequence were present in the flanking regions of the poorly transcribed gene copies \( tRNA^{\text{gus}} \)-6 and -7 [16,17]. Mutations in, or deletions of, the TATAAA sequence increased the transcriptional levels in vitro substantially. The fact that the TATAAA motif, a perfect polymerase II (pol II) consensus sequence for TBP binding, could exert a negative effect on pol III transcription was somewhat intriguing, since TBP is also required for pol III transcription [32]. These multiple TATAAA elements possibly competed for transcription factors such as TBP or the TBP-associated factors (TAFs), and brought about inhibition by sequestration of these factors.

In this study, we report the purification and characterization of a novel 43 kDa protein, designated here as P43 TBF (TATA-box-binding factor) from the posterior silk gland nuclear extracts, specifically binding to the TATAAA sequences. Supplementation of P43 TBF to the in vitro transcription reactions showed a pronounced inhibitory effect on transcription by both pol III and pol II.

**EXPERIMENTAL**

**Plasmid constructs**

Individual members of the \( tRNA^{\text{gus}} \) multi-gene family from *B. mori*, available from our laboratory stocks, were utilized. The gene copies used were \( tRNA^{\text{gus}}-1 \) and -11 (highly transcribed), \( tRNA^{\text{gus}}-2 \) and -4 (medium-level transcription) and \( tRNA^{\text{gus}}-6 \), -7 and -10 (poorly transcribed) [16,17]. *B. mori* \( tRNA^{\text{gus}} \) gene was a 420 bp genomic fragment in pBR327, encoding the 72 nt mature \( tRNA^{\text{gus}} \) [33]. The plasmid construct pAc5.1 OPEN (open reading frame) 42 containing ORF42 from *B. mori* nucleopolyhedrovirus under the control of the *Drosophila* actin promoter (constructed in our laboratory) was used as a template for pol II-mediated transcription. The cDNA clone of *B. mori* TBP carrying the 1.1 kb insert in the expression vector pET 19b [34] was a gift from Dr. Karen Sprague (University of Oregon, Eugene, OR, U.S.A.).

**In vitro transcription assays**

Nuclear extracts from posterior silk glands of *B. mori* in the fifth larval instar were prepared as described previously [13]. Briefly, freshly dissected silk glands (or glands frozen at \(-70^\circ\)C for up to 6 months) were homogenized (Dounce homogenizer) in 10 mM Hepes, pH 7.9, containing 2 M sucrose, 10\%, (v/v) glycerol, 15 mM KCl, 0.5 mM dithiothreitol (DTT), 0.5 mM PMSE, 0.15 mM spermine, 0.15 mM spermidine and 1 mM EDTA. Nuclei were pelleted by centrifugation (25000 rev./min in an SW 28 rotor for 70 min at 4 \( ^\circ\)C) and lysed in 20 mM Hepes, pH 7.9, containing 25\% (v/v) glycerol, 0.42 M NaCl, 1.5 mM MgCl\(_2\), 0.2 mM EDTA, 0.5 mM DTT and 0.5 mM PMSE. The crude lysate was cleared by centrifugation, dialysed against buffer A (20 mM Hepes, pH 7.9, containing 20\% glycerol, 0.1 M KCl, 0.2 mM EDTA, 0.5 mM DTT and 0.5 mM PMSE), and used as crude nuclear extracts for transcription. In vitro transcription reactions in a final volume of 25 \( \mu l \) contained 20 mM Hepes, pH 7.9, 60 mM KCl, 6 mM MgCl\(_2\), 0.1 mM EDTA, 6 mM creatine phosphate, 50 \( \mu \)M each of ATP, CTP and UTP, 10 \( \mu \)M GTP, 5 \( \mu \)Ci of \( \beta\)-\[^{32}\]P\)GTP (3000 Ci/mmole), nuclear extract (20 \( \mu \)g of protein) and 4 \( \mu \)g/ml of the supercoiled plasmid DNA template. After incubation at 30 \( ^\circ\)C for 1 h, the reactions were terminated by the addition of 0.2\% (w/v) SDS and 10 mM EDTA. The samples were extracted once with phenol, and the transcripts in the aqueous layer were precipitated by the addition of 3 vols. of ethanol in the presence of 100 \( \mu \)g/ml glycogen (as a carrier). The precipitate was re-suspended in gel-loading buffer containing 80\% formamide, and subjected to electrophoresis on 7 M urea/8\% (w/v) acrylamide gels.

These nuclear extracts were also used for pol II transcription experiments from the actin5C promoter. The pol II transcription reaction contained, in a final volume of 25 \( \mu l \), 20 mM Hepes, pH 7.9, 60 mM KCl, 6 mM MgCl\(_2\), 0.1 mM EDTA, 6 mM creatine phosphate, 500 \( \mu \)M each of ATP, CTP, UTP and GTP, nuclear extract (20 \( \mu \)g of protein) and 8 \( \mu \)g/ml of the supercoiled plasmid DNA template. After incubation at 30 \( ^\circ\)C for 1 h, the reactions were terminated by the addition of 0.2\% SDS and 10 mM EDTA, and the transcripts were isolated as for pol III trancriptions. The pol II-generated transcripts were identified using primer-extension analysis for the Bm ORF42 transcript using \( \beta\)-\[^{32}\]P-end-labelled primer (ATTCGGAGGTGTTTCCGCC) and reverse transcriptase, in the presence of all four dNTPs. The primer-extended products were analysed on a 7 M urea/6\% (w/v) acrylamide gel by electrophoresis.

**Electrophoretic mobility shift assay (EMSA)**

DNA fragments harbouring the TATAAA sequence were derived either from the flanking regions of \( tRNA^{\text{gus}} \) copies or by the annealing of two complimentary oligonucleotides. The TATAAA sequence present upstream of \( tRNA^{\text{gus}}-1 \) was released as a 150 bp EcoRI–KpnI fragment (RK fragment). The TATAAA sequence present in the far upstream region in \( tRNA^{\text{gus}}-6 \) and -7 (at \(-871 nt with respect to +1 of \( tRNA^{\text{gus}}-6 \) was released as a 45 bp SacI fragment (Sac45 fragment) [16]. The actin promoter sequence containing the TATAAA element was released as a 300 bp NcoI–EcoRI fragment from the plasmid vector pAc5.1/V5 (obtained from Invitrogen). The synthetic oligonucleotides 1 (\( 5^{\prime}\)GATCGAACTTTGCTTATATAAAAATATACC\( 3^{\prime}\)) and 2 (\( 5^{\prime}\)GATCGGTATATTTTTATATAAGCAAAGTTC\( 3^{\prime}\)) were used in the EMSAs and in the preparation of affinity matrix were designed on the basis of the upstream sequence of \( tRNA^{\text{gus}}-1 \). The assay system for the EMSAs [35] contained, in a final volume of 15 \( \mu l \), 0.1–2 \( \mu \)g of protein (crude nuclear extract or purified fractions as indicated), 12 mM Hepes, pH 7.9, 40 mM KCl, 5 mM MgCl\(_2\), 4 mM Tris/HCl, pH 8.0, 0.6 mM EDTA, 0.6 mM DTT, 5\% (v/v) glycerol and 2 \( \mu \)g of double-stranded poly(dI–dC). After incubation at 4 \( ^\circ\)C for 15 min, the radiolabelled DNA fragment (10000 c.p.m.) harbouring the TATAAA sequence was added, and the binding reaction was allowed to proceed for a further 15 min. In retardation–chase experiments, varying amounts of the unlabelled specific competitors were also included. The binding reactions were terminated by adding 5 \( \mu l \) of the gel-loading buffer, and the samples were electrophoresed on 5\% polyacrylamide gels at 4 \( ^\circ\)C.

**Expression and purification of silkworm TBP**

*Escherichia coli* BL-2(DE3) cells transformed with the TBP cDNA construct were grown at 37 \( ^\circ\)C in Luria–Bertani broth containing 100 \( \mu \)g/ml ampicillin, and were induced by the addition of 500 \( \mu \)M isopropyl \( \beta\)-d-thiogalactoside during the mid-exponential phase of growth. After 2 h, the cells were harvested and lysed by sonication in 0.1 M NaCl/HMGN buffer [25 mM Hepes (pH 7.6)/12.5 mM MgCl\(_2\)/10\% glycerol/0.1\% Nonidet P40/4 mM 2-mercaptoethanol/0.1 mM PMSE]. The supernatant fraction was loaded on to a 5 ml heparin–Sepharose column (Pharmacia) equilibrated in 0.1 M NaCl/HMGN buffer. The column was washed with 25 ml of loading buffer, and the bound proteins were step-eluted with 0.4 M and 1 M NaCl/HMGN buffers. TBP was eluted with 1 M NaCl, and the peak fractions were pooled and loaded on to a 0.5 ml Ni\(^{2+}\)-Sepharose column (Novagen) pre-equilibrated with 0.7 M NaCl/HMGN.
A novel TATA binding factor

Figure 1 Purification of TATATA-sequence-specific-binding protein

(A) Schematic presentation of the purification strategy for the TATATA-sequence-binding protein from B. mori posterior silk gland nuclear extracts. (B) Individual fractions at each fractionation step were analysed for sequence-specific DNA binding using the 150 bp fragment (—150 to —300 nt upstream of the tRNAGly gene; see Figure 2A, top panel). The probe (RK fragment containing one TATATAA sequence) was isolated as the KpnI–EcoRI restriction fragment from the tRNAGly gene and was end-labelled using Klenow DNA polymerase in the presence of [γ-32P]dATP. For assay conditions of EMSA, see the Experimental section. Lanes were loaded as indicated on the figure. PC-B, Q-Sepharose and TBF refer to the fractions emerging from phosphocellulose, Q-Sepharose, and the protein eluted from affinity chromatography respectively. (C) Shows the electrophoretic pattern on SDS/PAGE of TBF (150 ng) eluted from the affinity matrix. The proteins were stained by silver nitrate [44]. A single band of 43 kDa is detected.

Buffer. The column was washed with binding buffer, and the bound proteins were eluted with 0.7 M NaCl/HMG buffer containing 500 mM imidazole, and dialysed against 0.1 M NaCl/HMG buffer for 5 h. The purity of TBP was confirmed by SDS/PAGE, and checked for its activity by restoration of tRNA<sup>gly</sup> transcription in nuclear extracts that were immuno-depleted of anti-TBP antibodies.

Purification of TBF

Phosphocellulose chromatography

The posterior silk gland nuclear extract was loaded on to a column of phosphocellulose (1 cm × 15 cm) in buffer A (see above). The column was washed with the loading buffer, and the bound proteins were eluted using a step gradient (0.35 M, 0.6 M
and 1 M) of KCl. Individual fractions were tested for the presence of TBF using EMSAs.

Q-Sepharose chromatography

The TBF-containing fraction (0.35 M KCl fraction) from the phosphocellulose column was dialysed against buffer A and loaded onto a Q-Sepharose column (2 ml) pre-equilibrated with buffer A at a flow rate of 3 column-vols./h. The column was washed with 5 column-vols. of buffer A, and the bound proteins were eluted using a 1 ml linear gradient of KCl (from 0.1–0.6 M) in buffer A. The fractions containing TBF were pooled, dialysed against buffer A for 4 h, and then processed for DNA-affinity chromatography.

DNA-affinity chromatography

The DNA-affinity matrix was prepared using the double-stranded oligonucleotides [generated by annealing of oligonucleotide 1 (conjugated to biotin at the 5′ end) and the complementary oligonucleotide 2]. The annealed oligonucleotides were coupled with 0.1 ml of streptavidin–agarose beads for 30 min at 20–22 °C. The efficiency of coupling was checked by estimating the amount of DNA in the input and the unbound fractions. Almost 100% coupling efficiency was achieved. The non-specific binding sites in the affinity matrix were blocked by incubation with three column-vols. of buffer A containing 0.25 mg BSA/ml for 30 min at 20–22 °C. The agarose beads (linked to the double-stranded oligonucleotides via the biotin–streptavidin interaction) were used as an affinity matrix for the purification of TBF.

The fraction containing TBF from the previous step was incubated initially with 50 μg of poly(dI-dC) (dI-dC)/ml for 30 min to bind all the non-specific DNA-binding proteins. This sample (final volume 0.2 ml) was incubated with the affinity matrix at 4 °C for 1 h with gentle swirling. The suspension was briefly centrifuged, and the matrix was washed with 10 ml of binding buffer. The bound proteins were eluted using buffer A containing 1 M KCl, dialysed against buffer A for 4 h, and then stored as aliquots at −70 °C.

Determination of the equilibrium constant for DNA binding

The equilibrium constants for specific binding reactions are provided by the following equations [36]:

\[
[CD] = [C^\circ] \cdot K_{eq} \cdot [D]/(1 + K_{eq} \cdot [D])
\]

(1)

and

\[
[CD]/[D] = -K_{eq} \cdot [CD] + K_{eq} \cdot [C^\circ]
\]

(2)

where [D] and [CD] are the free and bound species of the probe respectively, [C^\circ] denotes the total amount of DNA binding sites and \( K_{eq} \) is the equilibrium constant of the binding reaction. Eqn (1) describes a hyperbola, whereas eqn (2) is a straight-line plot, equivalent to a Scatchard plot. Standard DNA-binding reactions were performed with a constant amount of protein sample and various dilutions of labelled probe. The protein–DNA complex formed in the binding reaction and the unbound probe were quantified in a PhosphorImager apparatus following electrophoresis. The equilibrium constant was determined graphically from a plot of [CD]/[D] against [CD], according to eqn (2), where the negative reciprocal of the slope of the line yields the value of \( K_{eq} \) and the intercept on the y-axis gives the value of [C^\circ].

RESULTS

Purification of the factor binding to TATATAAA

We have reported previously that EMSAs with a DNA fragment derived from tRNA\(^{\text{Gly}}\)-1 (−150 to −300 nt upstream of the tRNA-coding region, designated as RK fragment and harbouring a single TATATAAA motif; see Figure 2A) [17] and nuclear extracts derived from posterior silk glands identified specific complexes. These labelled complexes could be ‘chased out’ in the presence of an excess of unlabelled competitor. We examined the binding of silkworm TBP (purified recombinant protein) to the RK fragment, but this did not reveal any stable complex formation under the experimental conditions used.

The binding factor was purified from the posterior silk gland nuclear extracts using the strategy depicted in Figure 1(A). The essential features of the purification scheme included the use of two conventional chromatographic columns (phosphocellulose and Q-Sepharose) and the use of a specific DNA-affinity matrix. In the first step, posterior silk gland nuclear extracts were fractionated on a phosphocellulose P11 column. The pol III transcription machinery separates into two fractions, one containing pol III in addition to the transcription factor TFIIIB (PC-B fraction), and the other containing TFIIIC (PC-C fraction) [37]. The PC-B fraction that showed binding to the probe was purified further by chromatography on a Q-Sepharose column. The eluate showing the maximum binding (the 0.38 M KCl fraction) was purified further on the oligonucleotide affinity matrix. For this step, the fractions from the Q-Sepharose column were incubated first with poly(dI-dC) · (dI-dC) to eliminate the non-specific DNA-binding proteins present in the fraction, and then applied to specific DNA matrix (double-stranded oligonucleotides linked to agarose beads via streptavidin–biotin interactions). The pre-treatment of the samples with non-specific DNA (poly (dI · dC)) before the specific oligonucleotide affinity column step improved the purity of the final samples. The bound proteins were eluted at high salt concentrations.

The purification profile of the TBF is shown in Figure 1(B). Although multiple complexes were detected with crude nuclear extract, only a single specific complex was detected from the phosphocellulose chromatography step onwards. SDS/PAGE analysis (Figure 1C) of the active fraction from the affinity column revealed a single polypeptide of 43 kDa. That the DNA-binding activity was associated with this purified band was confirmed by EMSAs using the protein eluted from the gel following SDS/PAGE (Figure 2B; lanes 8–10). The TATATAAA-binding activity of fractions generated during the course of purification is shown in Table 1. Total protein content was reduced from 70 mg in the crude nuclear extract to 4 μg in the purified fraction.

This purified protein was subjected to protein microsequence analysis, and partial amino acid sequences of the 12 peptides generated from this protein are presented in Table 2. Multiple peptide sequences did not match with any of the known TBFs, such as TBPs, TRF (TBP-related factor) or TLFs (TBP-like factors) from other sources. Extensive database searches did not reveal any matches to these peptides. This protein thus appears to be a novel TBF, which we have designated P43 TBF.

P43 TBF binds specifically to the TATATAAA sequence

To characterize further the purified P43 TBF and establish the target DNA specificity, TATATAAA sequence in the RK fragment was mutated to GATATCA (mutant available in the lab stocks [17]; Figure 2A). EMSAs performed with the parental, as well as the mutant, RK fragment (Figure 2B) revealed that P43 TBF
Figure 2 Specific binding of P43 TBF to the TATA box sequence

(A) The position of RK fragment (parental type or mutant) on the rRNA is shown. The sequence TATAATA is present at position 7–276 nt with respect to the tRNA-coding region. In the mutant RK fragment, the sequence TATAATA was mutated to GATATCA. (B) EMSA with parental and mutant RK fragments (Wt TATA and Mut TATA respectively) using purified P43 TBF, showing the increasing concentrations of P43 TBF that were used for binding to the end-labelled probes (10000 c.p.m.). Lanes: 1, free probe (wt RK); 2, binding in the presence of 100 ng of P43 TBF; 3–5, chase using 10, 100- and 1000-fold molar excess of unlabelled parental RK fragment; 6–8, chase using 100, 1000- and 10000-fold molar excess of mutant RK fragment.

Table 1 Quantification of TATATAA binding factor purifcation by DNA-binding activity

<table>
<thead>
<tr>
<th>Column</th>
<th>Total protein (mg)</th>
<th>Total activity (units)</th>
<th>Specific activity (units/mg)</th>
<th>Fold purification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclear extract</td>
<td>70</td>
<td>62500</td>
<td>892</td>
<td>–</td>
</tr>
<tr>
<td>Phosphocellulose</td>
<td>22.2</td>
<td>36038</td>
<td>1623</td>
<td>1.8</td>
</tr>
<tr>
<td>Q-Sepharose</td>
<td>0.72</td>
<td>12628</td>
<td>17538</td>
<td>19.6</td>
</tr>
<tr>
<td>DNA-affinity</td>
<td>0.004</td>
<td>4000</td>
<td>1000000</td>
<td>1121</td>
</tr>
</tbody>
</table>

Table 2 Sequences of peptides generated from protein microsequence analysis of P43 TBF

The peptide sequences shown below have been determined from two independent batches of the purified protein P43. The protein sequencing was carried out at the University of Virginia Biomedical Research Facility. Peptides were generated by tryptic digest of P43 TBF and the sequence of each peptide was analysed by MS. 'X' denotes either isoleucine or leucine; '?' designates a single unknown residue, and … designates an unspecified number of unknown residues.

<table>
<thead>
<tr>
<th>Peptide no.</th>
<th>Measured molecular mass (Da)</th>
<th>Peptide sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>678.6</td>
<td>VXSMTK</td>
</tr>
<tr>
<td>2</td>
<td>935.5</td>
<td>EAXTVFEK</td>
</tr>
<tr>
<td>3</td>
<td>1303.7</td>
<td>??EHLDHYR</td>
</tr>
<tr>
<td>4</td>
<td>1359.7</td>
<td>??PSDTV… K</td>
</tr>
<tr>
<td>5</td>
<td>1443.9</td>
<td>KGDOEDTGGFK</td>
</tr>
<tr>
<td>6</td>
<td>1478.7</td>
<td>FXDFESHXSRR</td>
</tr>
<tr>
<td>7</td>
<td>1536.7</td>
<td>??XAYSSVGEYK</td>
</tr>
<tr>
<td>8</td>
<td>1589.9</td>
<td>… EKAGEY… GK</td>
</tr>
<tr>
<td>9</td>
<td>1691.9</td>
<td>??SENXTIDVYK</td>
</tr>
<tr>
<td>10</td>
<td>1692.6</td>
<td>??XAYSSVGEYKR</td>
</tr>
<tr>
<td>11</td>
<td>2052.3</td>
<td>??MEFVEF??XADXDENK</td>
</tr>
</tbody>
</table>

Having established that the binding is specific to the TATATAA sequence, we also analysed the binding of P43 TBF to the TATATAA sequences present in the flanking region of rRNA and -7 [16], as well as the TATA sequence present in actin5C promoter, which is a typical pol II promoter. The position of the TATATAA sequence (harboured in the Sac45 fragment) with respect to the rRNA and -7 coding regions, and the location of the TATA box in the actin promoter, are shown in Figure 3(A). EMSAs performed with Sac45 fragment and the actin promoter revealed the presence of complexes proportional to the increasing amounts of P43 TBF added (Figure 3B). The specificity of binding to these fragments was also binding to the parental fragment was completely abolished when TATATAA was mutated to GATATCA (mutations are in bold). The binding-competition experiments performed using the parental TATATAA and the mutant sequences confirmed these results (Figure 2C). Even at a 1000-fold molar excess, when the parental sequence completely chased out the labelled complexes, the mutant sequence did not have any effect on the complex formed with the parental sequences.

Having established that the binding is specific to the TATATAA sequence, we also analysed the binding of P43 TBF to the TATATAA sequences present in the flanking region of rRNA and -7 [16], as well as the TATA sequence present in actin5C promoter, which is a typical pol II promoter. The position of the TATATAA sequence (harboured in the Sac45 fragment) with respect to the rRNA and -7 coding regions, and the location of the TATA box in the actin promoter, are shown in Figure 3(A). EMSAs performed with Sac45 fragment and the actin promoter revealed the presence of complexes proportional to the increasing amounts of P43 TBF added (Figure 3B). The specificity of binding to these fragments was also

(10 000 c.p.m.) with 10 ng of P43 TBF was performed, and the complex was 'chased' using increasing amounts of either unlabelled parental RK fragment or mutant RK fragment. Lanes: 1, free probe (wt RK); 2, binding in the presence of 100 ng of P43 TBF; 3–5, chase using 10, 100- and 1000-fold molar excess of unlabelled parental RK fragment; 6–8, chase using 100, 1000- and 10000-fold molar excess of mutant RK fragment.
confirmed by competition experiments in which they were used as competitors for binding to the labelled RK fragment. The P43 TBF–RK fragment complex was competed out with increasing concentrations of either the Sac45 or actin promoter fragments with equal efficiency, reaching almost complete competition at 1000-fold molar excess of the competitor (Figure 3C). P43 TBF therefore appears to be a general TBF involved in the regulation of transcription.

Equilibrium constant of P43 TBF binding to TATATAA sequences

EMSA, which separates DNA probe–protein complex from unbound probe, provides a simple assay for studying the kinetics of specific DNA–protein interactions [36]. Figure 4(A) shows the result of a titration experiment in which P43 TBF complex formation was measured as a function of increasing concentrations of the DNA probe. Since the P43 TBF used for the binding reaction was pure, addition of excess carrier DNA to quench the non-specific binding of other proteins was not required. Therefore the binding reactions did not contain any DNA other than the specific probe. The equilibrium constant $K_{eq}$ for P43 TBF binding determined from the quantification of binding data from Figure 4(A) was $4.0 \times 10^{11} \text{M}^{-1}$, and the total number of binding sites $[C']$ was $3.04 \times 10^{-8} \text{M}$ (Figure 4B).

**Figure 3  Binding of P43 TBF to TATA motifs present in other genes**

(A) The position of Sac45 fragment harbouring the TATATAA sequence on the tRNA$^{Gly}$-5,7 construct is shown. The RNA-coding region is marked with an arrow. The location and sequence of actin promoter TATA box in construct pAc5.1 V5 is shown. (B) EMSAs with the Sac45 and RK (150 bp) fragments, and the 300 bp actin promoter region (Act TATA). All three fragments were end-filled in the presence of $[\alpha-\text{32P}]\text{dATP}$, and were used as probes for EMSAs (10000 c.p.m. of each probe). Lanes: 1–3, binding with Sac45 probe; 4–6, binding with RK probe; 7–9, binding with actin TATA box as probe. The positions of the complex are indicated by arrows. (C) The complex between the wild-type RK fragment and P43 TBF (10 ng) was chased with excess unlabelled Sac45 fragment or unlabelled actin TATA fragment. Lanes: 1, free probe (labelled RK fragment); 2, binding in the presence of purified P43; 3–5, chase using 10-, 100- and 1000-fold molar excess of unlabelled Sac45 fragment; 6–8, chase using 10-, 100- and 1000-fold molar excess of unlabelled actin promoter fragment.

**P43 TBF inhibits in vitro transcription from both pol II and pol III promoters**

The effect of P43 TBF on *in vitro* pol III transcription of tRNA$^{Gly}$ gene copies was analysed by supplementing the *in vitro* transcription assays using posterior silk gland nuclear extracts with P43 TBF. Figure 5(A) shows the effect of P43 TBF on the transcription of tRNA$^{Gly}$-1. Transcription from tRNA$^{Gly}$-1 gives rise to two distinct transcripts, corresponding to an approx. 81 nt primary transcript (precursor) and a 71 nt mature (processed) transcript. With increasing concentrations of P43 TBF, transcription was decreased markedly, and was completely inhibited in the presence of 40 nM P43 TBF (Figures 5A and 5B). The inhibition of transcription was consistent, and the levels of inhibition were identical with different batches of nuclear extract and purified P43 TBF. The inhibition of the transcription due to P43 TBF was also confirmed using the protein eluted from the gel following SDS/PAGE (Figure 5A, lanes 6–8).

The effect of P43 TBF on the transcription of other tRNA$^{Gly}$ gene copies was also analysed (Figure 5C). The gene copies examined belonged to the three different categories of expression, namely tRNA$^{Gly}$-1 and -11 (highly transcribed), tRNA$^{Gly}$-2 and -4 (medium-level transcription) and tRNA$^{Gly}$-6, 7 and -10 (poorly transcribed). Supplementation of 30 nM of the purified protein
resulted in complete inhibition of in vitro transcription of all the tRNA<sup>Ala</sup>-1 genes tested.

The effect of P43 TBF on the in vitro transcription of a tRNA<sup>Ala</sup>-1 construct, where the TATATAA box sequences have been deleted, as well as the transcription of a different tRNA gene, tRNA<sup>Glu</sup>-1 from B. mori, was analysed. On supplementation with increasing concentrations of P43 TBF, the transcription of both the genes were markedly decreased (Figure 5D). The quantification of transcription is presented in Figure 5E). As we reported previously [17], transcription from the tRNA<sup>Ala</sup>-1 construct lacking the TATATAA sequences was found to be 20–50 % greater than that of the parental construct. The transcription from both these constructs were equally susceptible to inhibition by the externally added P43 TBF, and more than 90 % inhibition was seen at a concentration of 40 nM protein.

Having established that in vitro supplementation of P43 TBF to crude nuclear extracts causes inhibition of tRNA gene transcription, the effect of this protein on pol II transcription was analysed. The posterior silk gland nuclear extracts from B. mori showed appreciable levels of pol II transcription from the Drosophila actin<sup>5C</sup> promoter. The amount of nuclear protein required for maximum transcription from the actin promoter was standardized. Maximal transcriptional activity was achieved in the presence of 15 μg of nuclear protein, and a specific transcript of approx. 350 nt, corresponding to the inserted fragment (Bm ORF42), under the expression of the actin promoter was obtained. The effect of P43 TBF on transcription from the actin promoter is shown in Figure 6(A). Addition of 40 nM P43 TBF caused approx. 60 % inhibition of transcription from the actin promoter (Figure 6B). To confirm that the transcript arising from the actin promoter was indeed due to pol II, we also analysed the sensitivity of the transcription to α-amanitin. Among eukaryotic polymerases, pol II is most sensitive to α-amanitin (50 % inhibition at 25 ng/ml) and pol III shows an intermediate level of sensitivity (50 % inhibition at 20 μg/ml), whereas pol I is resistant. Addition of 50 ng/ml of α-amanitin to in vitro transcription reactions completely abolished the transcription from the actin promoter (Figure 6A, lanes 7–10), indicating that the transcription from the actin promoter was indeed due to pol II.

P43 TBF inhibited both pol II and pol III transcription; however, the inhibition was more pronounced in the case of pol III. TBP is a common factor involved in transcription by both polymerases, being a component of TFIIIB and TFIIID. The nuclear extracts, when supplemented with increasing concentrations of recombinant TBP, caused a decrease in the in vitro transcription levels of tRNA<sup>Ala</sup>-1, and the transcript was almost undetectable after supplementation with excess TBP (Figure 7A, lanes 1–6). The sequestration of TAFs is a possible cause for this inhibition by excess TBP. The nuclear extracts, when immunodepleted of TBP with TBP-specific antibodies, showed a marked reduction in the levels of tRNA<sup>Ala</sup>-1 transcription (Figure 7A, lanes 7 and 8). The transcription by these immunodepleted extracts was restored upon addition of purified recombinant silkworm TBP (lanes 9–12). Evidently, TBP was essential for transcription. In order to see if inhibition caused by P43 TBF was due to direct competition with TBP, we examined whether externally supplemented TBP could reverse the inhibitory effect of P43 TBF (Figure 7B). Supplementation of TBP did not reverse the inhibition caused by P43 TBF. On the contrary, addition of TBP augmented the inhibitory effect of P43 TBF (lanes 4–6). We also analysed the effect of P43 TBF on transcription of tRNA<sup>Glu</sup>-1 with nuclear extracts immunodepleted of TBP antibodies (Figure 7C). Transcription by the immunodepleted extracts was not restored by supplementation with purified P43 TBF (lanes 3–5). However, as expected, recombinant TBP restored the transcription of tRNA<sup>Glu</sup>-1 (lanes 6 and 7).

These results support the hypothesis that both TBP and P43 TBF can both contribute to such regulation of pol III transcription; however, tissues such as the silk-glands of B. mori might provide exceptions to this rule, where a functional adaptation as reflected by enhanced synthesis of certain species of tRNA (e.g. tRNA<sup>Glu</sup> and tRNA<sup>Ala</sup>) takes place to optimize the committed synthesis of the silk proteins [27,28]. In such instances, more regulatory circuits might be operative to achieve preferential and enhanced transcription of specific tRNA genes, as demanded by tissues with a committed function. The presence of cis-regulatory elements, and the trans-acting factors capable of interacting with them, can both contribute to such regulation of pol III transcription of tRNA genes.

The different copies of tRNA<sup>Glu</sup> in B. mori are differentially transcribed both in vitro and in vivo, and hence form a good model system for the analysis of tRNA gene regulation. The
Figure 5  Effect of P43 TBF on pol III transcription

P43 TBF was supplemented to in vitro transcription reactions. For details of the in vitro transcription assays with posterior silk gland nuclear extracts, see the Experimental section. Various amounts of P43 TBF were added to the nuclear extracts, as indicated, and incubated for 10 min in the presence of transcription buffer and NTP mix. To initiate the transcription reaction, 100 ng of the plasmid template was added and the samples were incubated for 60 min at 30°C. (A) Effect of supplementation of increasing amounts of P43 TBF to the in vitro transcription of \( tRNA^{\text{Gly}} \). Distinct transcripts, corresponding to the precursor tRNA (81 nt) and processed tRNA (71 nt), are shown. In vitro transcription of \( tRNA^{\text{Gly}} \) in the absence of supplemented P43 TBF (lane 1), and in the presence of 10, 20, 30 and 40 nM P43 TBF (lanes 2–5 respectively) is presented. The panel to the right (lanes 6–8) shows the inhibition of pol III transcription \( tRNA^{\text{Gly}} \) by P43 TBF eluted out.
upstream sequences of \( tRNA^{Gly} \) copies, located up to −100 nt with respect to the \( tRNA \)-coding region, are highly A/T-rich [14]. Transcription of the different \( tRNA^{Gly} \) copies is negatively modulated by the presence of the TATATAA sequence in the far-flanking regions of the coding region [13,16,17]. However, similar sequences (with changes in one or two nucleotides) were also present within the first 100 nt in 5′ flanking regions of many \( tRNA^{Gly} \) copies. For instance, in addition to the far-flanking TATATAA sequence, which exerted the negative effect on transcription, \( tRNA^{Gly}_{1-1} \) and \( tRNA^{Gly}_{1-6} \) harboured perfect TATAATAA sequences at nt positions −86 and −26 respectively. These ‘near-TATATAA’ elements, when deleted, resulted in a loss of transcription [13,17]. Moreover, the distal TATATAA sequence, when moved closer to the transcription start site, acted as a positive modulator for transcription in the absence of other positive regulatory elements [17].

Although most regions upstream of the \( tRNA \) gene do not harbour TATA-box sequences, TBP is an essential component of the pol III transcription machinery, as in the case of pol II and pol I transcription [38]. The recruitment of TBP to the pol III transcription initiation sites is generally achieved as a component of the TFIIIB, by interaction with TFIIIC. Some of the pol III-transcribed genes, such as that for the U6 snRNA, however, harbour TATA sequences in the upstream region and are known to recruit TBP directly via interaction with these elements [1].

We have purified a novel TBP, P43 TBF, from the posterior silk gland nuclear extracts of \( B. \) \( mori \). Microsequence analysis of a few peptides generated from this protein did not show matches to any in the database entries. It is now well established that, besides TBP, metazoans also encode homologues of TBP, designated as TRFs. Two classes of TRF have been identified so far [39]. The shared homology of TBP and TRFs is confined to their DNA-binding core repeat domains. TRF1, reported only from \( Drosophila \) (dTRF1), is involved in both pol II and pol III transcription and is able to bind to TATA boxes present in pol II-associated promoters [40,41]. It also plays a major role in pol III transcription in \( Drosophila \), and is found to be associated with the \( Drosophila \) TFIIIB [41]. In contrast, TRF2s, also termed

---

**Figure 6** Effect of P43 TBF on pol II transcription

The effect of P43 TBF on transcription by pol II from the \( Drosophila \) actin promoter was analysed in \textit{vitro} using the posterior silk gland nuclear extracts. The transcripts generated were detected by primer-extension analysis using a primer specific for the reporter gene (ORF42 from \( BmNPV \) cloned under the actin promoter), and situated 350 nt from the transcription start site. (A) The effect of increasing concentrations of P43 TBF on transcription from the actin promoter. Lanes: M, molecular-size marker (pTZ 18 DNA digested with \textit{Hind} I and end-labelled); 1, control (no template DNA); 2 and 6, \textit{in vitro} transcription from actin template in the absence of supplemented P43 TBF; and 3, 4 and 5, in the presence of 20, 10 and 40 nM of P43 TBF respectively; 7–10, \textit{in vitro} transcription from actin template, in the absence (lane 7) or presence (lanes 8–10) of 10, 25 or 50 ng/ml \( \alpha \)-amanitin. Quantification of transcription (by PhosphorImaging) is shown in (B).
The binding of P43 TBF to the TATATAA boxes was highly sequence-specific, and mutation of the sequence to GATATCA abrogated the binding. However, the protein was able to bind to the TATA box (TATAAA) sequence of a canonical pol II promoter (Drosophila act5C). Thus P43 TBF could be a general TATA-binding factor. Analysis of the equilibrium constant for the binding of P43 TBF to a specific DNA sequence revealed it to be a much stronger DNA-binding protein ($K_{eq}$ = $4.0 \times 10^{10}$ M$^{-1}$) than silkworm TBP or yeast TBP, for which $K_{eq}$ values of 2.4 $\times 10^{9}$ M$^{-1}$ and 3 $\times 10^{9}$ M$^{-1}$ respectively were determined for binding to the adenovirus major late promoter TATA sequence [34,43].

Supplementation of P43 TBF inhibited both pol II and pol III transcription in vitro. It was possible that the protein competed with the TBP-binding sequences and brought about inhibition. However, the inhibition by P43 TBF was independent of the presence of the cognate DNA-binding site. For instance, transcription from tRNA$^{\text{Gly}}$-1 construct, in which the TATATAA sequence was deleted, was inhibited to the same extent as the parental copy. Moreover, transcriptional inhibition by P43 TBF was not reversed on addition of TBP. On the other hand, there was a further decrease in transcription by externally supplemented TBP. In fact, externally supplemented TBP by itself brought about inhibition of transcription, and evidently was not limiting in crude nuclear extracts. However, removal of TBP by antibody treatment abolished pol III transcription, and the inhibition could be reversed by externally added TBP, but not by P43 TBF. In the absence of the P43 TBF clone, we could not generate antibodies against the protein to analyse the direct effect of its deletion from the nuclear extracts, on transcription.

From the above results, it is clear that P43 TBF, in addition to binding to the cognate sequence elements, inhibited transcription by interacting with some other components of the transcription machinery. It is likely that some other transcription factor that associates with TBP (e.g., the subunits of TFIIIB or some of the TAFs necessary for pol II/pol III transcription) become limiting due to sequestration by the externally supplemented P43 TBF (or TBP), resulting in inhibition of transcription. We conclude that when P43 TBF is bound at a distal region from the transcription start site, or when available in excess, this sequesters some essential transcription factor.

We thank S. Srikanth and K.S. Ponnoovil, Central Silk Research and Training Institute, Central Silk Board, Mysore, India, for the generous supply of silkworm larvae, Karen Sprague for the silkworm TBP cDNA clone and Asha Acharya (from our laboratory) for the plasmid construct pAcV5 Bm ORF42. We are grateful to the Divisional facility for automated DNA sequencing and phosphorimager use. We also thank the Department of Science and Technology, Government of India for financial support. Protein sequencing was carried out at the W.M. Keck Biomedical Mass Spectrometry Laboratory of the University of Virginia Biomedical Research Facility, VA, U.S.A. L.S. is a recipient of a research fellowship from the Council for Scientific and Industrial Research.

REFERENCES


21 Carbon, P. and Krol, A. (1991) Transcription of the Xenopus laevis selenocysteine tRNA(SeCys) gene. A system that combines an internal B-box and upstream elements also found in U6 snRNA genes. EMBO J. 10, 599–606


Received 2 November 2001/7 January 2002; accepted 12 February 2002

© 2002 Biochemical Society