An inverted TATA box directs downstream transcription of the bone sialoprotein gene

Jack J. Li,* Richard H. KIM† and Jaro SODEK†‡

MRC Group in Periodontal Physiology, *Faculty of Dentistry and †Department of Biochemistry, University of Toronto, Toronto, Ontario, Canada M5S 1A8

The orientation of the TATA box is thought to direct downstream transcription of eukaryotic genes by RNA polymerase II. However, the putative TATA box in the promoter of the bone sialoprotein (BSP) gene, which codes for a tissue-specific and developmentally regulated bone matrix protein, is inverted (5'-TTTATA-3') relative to the consensus TATA box sequence (5'-TATAAA-3') and is overlapped by a vitamin D₃-response element. Here we show that the inverted TATA sequence in the rat BSP gene binds to recombinant TATA-box-binding protein (TBP) with an affinity similar to that observed with the consensus TATA box, and site-directed point mutations in the inverted TATA sequence (mutating TTTATA into TCTCTA) abrogate both TBP binding and BSP promoter activity. However, when

the inverted TATA sequence is changed to a canonical TATAAA, the TBP- and vitamin D₃ receptor-binding properties together with the BSP promoter activity are retained. In addition, we found that the TBP is required to reconstitute *in vitro* transcription driven by the BSP promoter. These studies, which have revealed a naturally occurring inverted TATA box that can bind TBP and direct downstream transcription, demonstrate that the orientation of the TATA box does not determine the direction of transcription in higher eukaryotic genes. Consequently, the inverted TATA box that is conserved in the human, rat and mouse BSP gene promoters will provide an excellent *in vivo* model to investigate the polarity of the transcription factor IID–DNA complex and its relation to downstream transcription.

INTRODUCTION

The transcription of eukaryotic genes is highly regulated by interactions between a variety of *trans*-acting factors and their cognate *cis*-acting elements. The binding of the transcription factor (TF) IID complex to the TATA box is crucial for the initiation of transcription and for controlling gene expression (reviewed in [1,2]). Even in genes lacking a recognizable TATA box, binding of TFIID to a region approximately –30 bp upstream from the transcription start site is required to initiate transcription by RNA polymerase II ([3] and references therein, [4]). Included in the TFIID complex is the universal transcription factor TATA-box-binding protein (TBP), which specifically recognizes the TATA box, and the TBP-associated factors (TAFs) [3,5,6]. Thus the TATA box provides a specific site that directs TFIID, RNA polymerase II and other auxiliary factors to assemble the pre-initiation complex (PIC) which promotes the synthesis of mRNA at the cap site, and mediates the activity of upstream activator elements [2,7–13].

Since most TATA sequences found so far are asymmetrical (5'-TATAAA-3' as a canonical form) [7,14,15], the TATA box is thought to be capable of orienting TBP and other PIC factors thereby defining the direction of transcription [1,15–17]. In support of this contention, it has been shown that conversion of a consensus TATA box into an inverted sequence (5'-TTTATA-3') is associated with a loss of downstream gene transcription activity [18,19]. However, some experiments in which the TATA sequence was inverted artificially indicated that the orientation of the TATA box might not determine the direction of transcription [20,21]. Consequently, the importance of the TATA box orientation in the determination of the TBP/TATA polarity and the direction of transcription is still ambiguous [20–22].

Bone sialoprotein (BSP) has been characterized as a major non-collagenous protein, mainly found in the extracellular matrices of bone, cementum and dentine, the expression of which is developmentally regulated ([23,24] and references therein). In addition to its specific expression by fully differentiated osteoblasts at sites of *de novo* bone formation, BSP is able to nucleate hydroxyapatite crystal formation from steady-state physiological concentrations of Ca²⁺ and P_i [25]. Therefore BSP is believed to function in the initial formation of the mineralized bone tissue [24]. To study the transcriptional regulation of BSP gene expression, we have recently cloned the rat BSP gene promoter [26]. Intriguingly, a perfect inverted TATA sequence (5'-TTTATA-3') is located at –24 to –19 bp upstream of the transcription initiation site in the rat BSP gene promoter. Subsequent studies in our laboratories [27], as well as others [28], have shown that this inverted TATA sequence is conserved in the human and mouse (A. Gupta and J. E. Aubin, unpublished work) BSP gene promoters, implying its importance in controlling BSP gene expression. Here, we show that the inverted TATA element in the rat BSP promoter is a functional TATA box that is required for driving downstream expression of this eukaryotic gene both *in vitro* and *in vivo*. This study provides the first characterization of a naturally occurring inverted TATA box in eukaryotic genes and demonstrates that the specific orientation of the TATA box does not determine the direction of transcription. As a vitamin D-response element (VDRE) overlaps this inverted TATA box and mediates the suppression of BSP gene transcription by 1,25-dihydroxyvitamin D₃ (R. H. Kim, J. J. Li, Q. Zhang, Y. Ogata, M. Yamauchi, L. P. Freedman and J. Sodek, unpublished work), it was also of interest to determine whether the inverted TATA box was required to accommodate a functional VDRE.

Abbreviations used: BSP, bone sialoprotein; CAT, chloramphenicol acetyltransferase; CRE, cyclic AMP-response element; PIC, pre-initiation complex; TBP, TATA-box-binding protein; TAF, TBP-associated factor; VDR, vitamin D₃ receptor; VDRE, vitamin D-response element; SV40, Simian virus 40; TF, transcription factor.

‡ To whom correspondence should be addressed.

MATERIALS AND METHODS

Comparison of the rat and human BSP gene promoters

Part of the cloned rat BSP gene promoter was analysed in the Transcription Factor Database using the University of Wisconsin Genetics Computer Group (GCG) program. The nucleotide sequences of the rat and human BSP gene promoters were then compared for homology and aligned by the BESTFIT command of the GCG program.

DNase footprinting

To prepare DNA probes for DNase I footprinting experiments, plasmid pCAT3 [26] was digested with either *HindIII* or *SalI*. Either the *HindIII* end (for the coding strand) or the *SalI* end (for the non-coding strand) was first labelled with [γ - 32 P]ATP (3000 Ci/mmol) by T4 polynucleotide kinase. The end-labelled 176 bp DNA fragments (position -116 to +60 of the rat BSP promoter; see Figure 1) were then digested with secondary restriction enzymes, *SalI* or *HindIII*, and purified by PAGE. A specific radioactivity of approx. 2×10^7 c.p.m./ μ g was usually obtained. About 5 fmol (approx. 0.5 ng) of 32 P-labelled DNA probe was used in each footprinting reaction. Footprinting experiments were performed as described [29], with slight modification, in a reaction mixture (50 μ l) containing 12 mM Tris/HCl (pH 7.9), 40 mM Hepes (pH 8.4), 60 mM KCl, 12% (v/v) glycerol and 8 mM MgCl₂. After incubation of end-labelled DNA probes with various amounts of protein at 25 °C for 45 min, DNase I digestions were carried out at room temperature (22 °C) for 2 min. Products were analysed on a urea/10% polyacrylamide sequencing gel. TBP-DNA binding affinity was determined as described by Ausubel et al. [30]. The absorbance over footprint sites was digitized and quantified by two-dimensional densitometry of autoradiograms using a Macintosh Apple One Scanner system. Dissociation constants (K_d) were calculated from data calibrated to internal standards and expressed using the Langmuir isotherm.

Site-directed mutagenesis

Site-directed point mutations were generated by PCR by the method of Landt et al. [31] with modifications. Briefly, two separate reactions were performed in the first round of amplification. One reaction mixture contained 1 ng of plasmid pCAT3 mixed with a 5'-flanking primer, antisense mutant primer, plus the necessary PCR components. Another reaction mixture contained identical components except for a 3'-flanking primer and the sense mutant primer. After 20 cycles of amplification, one mutant fragment with the mutation at the 3' end and another mutant fragment with the mutation at the homologous 5' end were generated. These two PCR products were then separated on a low-melting-point agarose gel, sliced, combined and melted. A small volume (10 μ l from about 200 μ l) was taken to perform the second-round PCR with the 5'-flanking primer and the 3'-flanking primer used in the first-round PCR. The final PCR fragment containing the desired point mutation was subcloned into pCAT-Basic vector. The designated mutations and fidelity of PCR were confirmed by dideoxy sequencing (see Figure 2).

Chloramphenicol acetyltransferase (CAT) plasmid construction and CAT assays

Transient transfections and CAT assays were conducted as described previously [26]. The transient expression vector pCAT-Basic (Promega) containing the bacterial CAT reporter gene, but lacking a promoter and enhancers, was used to test for BSP gene

promoter activity. The wild-type and mutated rat BSP promoters (-116 to +60) were cloned into the *HindIII*-*SalI* sites of the pCAT-Basic vector. The osteoblastic cell line, ROS17/2.8, was grown in α -minimal essential medium containing 10% (v/v) fetal bovine serum and seeded 24 h before transfection at a density of 4×10^5 cells per 100 mm dish. Transfections were performed by the calcium phosphate precipitation technique [30], using 10 μ g of BSP-CAT fusion plasmid DNA for each dish. Normalization of CAT enzyme activity for transfection efficiency was determined by co-transfection with 3 μ g of pSV- β -galactosidase control vector (Promega) containing the *lacZ* gene driven by the Simian virus 40 (SV40) promoter. Cells were harvested at 48 h after transfection, and cell extracts were prepared by repeated freeze-thawing. Liquid-scintillation counting was employed to determine CAT activity (Promega's instructions).

In vitro transcription

The templates for the *in vitro* transcription reaction were prepared by PCR based on wild-type or mutated BSP-CAT fusion plasmids described above. One 5'-primer was derived from the rat BSP promoter (-116 to -107) and two 3'-primers were derived from sequences within the CAT reporter gene to produce templates that could generate transcripts of 171 bp or 311 bp in size. The reaction was carried out as previously described [32,33] with slight modifications. Briefly, each reaction was performed in 25 μ l of 8.8 mM Hepes, pH 7.9 at 25 °C, containing 8.8% glycerol, 44 mM KCl, 0.088 mM EDTA, 0.22 mM dithiothreitol, 10 mM MgCl₂ and 200 ng (8 mg/ml) of DNA template. The reaction was started by mixing the HeLa-cell nuclear extract (Promega) and DNA and incubating for 15 min at 30 °C. Transcription was initiated by adding ATP, CTP and UTP (200 mM each), 16 mM GTP and 10 μ Ci of [α - 32 P]GTP (3000 Ci/mmol) and incubating at 30 °C for a further 45 min. The reaction was terminated by adding 175 μ l of 0.3 M Tris/HCl, pH 7.5 at 25 °C, 0.3 M sodium acetate, 0.5% SDS, 2 mM EDTA and 3 mg/ml tRNA. RNA was extracted with phenol/chloroform and precipitated with ethanol. The pellet was washed once with 70% ethanol and dissolved in 10 μ l of loading dye; each reaction mixture was analysed on a urea/5% polyacrylamide denaturing gel. In the reconstitution assay, the HeLa-cell nuclear extract was heated at 45 °C for 15 min to inactivate the endogenous TBP [34]. To this heat-treated nuclear extract was added either BSA or various amounts of TBP, and the *in vitro* transcription reactions were carried out as described above.

RESULTS

Sequence comparison between the rat and human BSP gene promoter

The rat and human BSP promoters have been cloned and described in previous studies and upstream DNA sequences of these promoters shown to be functional in directing the bacterial CAT reporter gene expression in osteoblastic cells [26,27]. Analysis of the 5'-flanking sequences (+1 to -801) of the rat BSP gene revealed a number of sequences resembling consensus eukaryotic *cis*-acting elements [27]: an inverted CCAAT box (ATTGG) at -50 to -46, a nuclear factor I (NFI)-binding site (GATTGGCT) at -51 to -44, a cyclic AMP-response element (CRE) (TGACGTCG) at -75 to -66, an NF κ B-binding site (CGGATTTTCT) at -102 to -93, an AP1-binding site (TTATTCA) at -143 to -137, a homeobox-binding site (TCAATTAAAT) at -194 to -185, an AP2-binding site (GGCAGCCC) at -447 to -440 and a novel 18-nucleotide

Figure 1 Nucleotide sequence of the immediate promoter region of the BSP gene encompassing the inverted CCAAT and TATA boxes

The region protected by human TBP in the DNase I footprinting experiments, whereby human TBP is shown to protect an approx. 20 bp region encompassing the inverted TATA box (see Figure 2). The putative VDRE overlapping the inverted TATA box is shadowed.

Figure 2 DNase I footprinting experiments: binding of recombinant human TBP to the inverted TATA element in the rat BSP gene promoter

(a) Footprint of human TBP on the coding strand; (b) footprint of human TBP on the non-coding strand. DNA co-ordinates including the position of the inverted TATA box are included in (a) and (b). The nucleotide position was deduced from G + A sequencing reactions [lane 1 in (a) and (b)] (M, markers). BSA was used as a control [lane 2 in (a) and (b)]. Labelled DNA probes were incubated with increasing amounts of human TBP. Amounts of human TBP used in lanes 3–7 in (a): 1/16 footprinting units (FPU) (= 1.25 ng), 1/8 FPU (= 2.5 ng), 1/4 FPU (= 5 ng), 1/2 FPU (= 10 ng), 1 FPU (= 20 ng). Amounts of human TBP used in lanes 3–9 in (b): 1/64 FPU (= 0.3125 ng), 1/32 FPU (= 0.625 ng), 1/16 FPU (= 1.25 ng), 1/8 FPU (= 2.5 ng), 1/4 FPU (= 5 ng), 1/2 FPU (= 10 ng), 1 FPU (= 20 ng).

Figure 3 Effect of point mutations on the binding of human TBP to the inverted TATA box

Recombinant human TBP is shown to bind to the wild-type inverted TATA sequence (TTTATA in pCAT3) (lanes 2–7), as well as to the canonical TATA sequence (TATAAA in p3NT-CAT) which was created by mutagenesis (lanes 14–19) with similar binding affinity. Two point mutations in the inverted TATA element (TCTCTA in p3ΔT-CAT) abrogated the TBP–DNA interaction (lanes 8–13). The nucleotide position was deduced from G + A sequencing reactions (lane 1; M, markers). Labelled DNA probes were incubated with BSA controls (lanes 2, 8, 14) and increasing amounts of human TBP: 1/16 FPU (= 1.25 ng) (lanes 3, 9, 15); 1/8 FPU (= 2.5 ng) (lanes 4, 10, 16); 1/4 FPU (= 5 ng) (lanes 5, 11, 17); 1/2 FPU (= 10 ng) (lanes 6, 12, 18); 1 FPU (= 20 ng) (lanes 7, 13, 19).

inverted repeat at –453 to –436. Notably, an inverted TATA sequence (5′-TTTATA-3′) in the region expected for a TATA box was identified at –24 to –19 (Figure 1) and an identical sequence observed in the human [27] and mouse (A. Gupta and J. E. Aubin, unpublished work) BSP gene promoters. Moreover, this inverted TATA sequence is overlapped by a VDRE (R. H. Kim, J. J. Li, Q. Zhang, Y. Ogata, M. Yamauchi, L. P. Freedman and J. Sodek, unpublished work). Comparison of the promoter and the first exon of the rat BSP gene with the corresponding region of the human BSP gene revealed a high degree of sequence identity with 71 % of the nucleotides conserved. The sequence identity is even higher in the region proximal to the transcription start site, with the inverted CCAAT box, NFκB-binding site, CRE, homeobox-binding site and an API-

binding site, as well as the inverted TATA box, being conserved in the rat and human BSP gene promoters.

TBP binds to the inverted TATA box

As the binding of TBP to the TATA box is a preliminary step in the formation of the PIC that is required for basal transcription of a class-II gene promoter, we used DNase I footprinting to assess the ability of TBP to bind to the inverted TATA element. Recombinant human TBP was demonstrated to bind to an approx. 20 bp region covering the inverted TATA element on both DNA strands (Figures 2a and 2b), similarly to its interaction with consensus TATA box sequences [14,34–37]. Moreover, unlike many other non-consensus TATA elements that interact

Figure 4 Effect of point mutations on the transcriptional activity of the rat BSP promoter

The bar values represent means $\pm$ S.E.M. from ten independent transfection experiments, each performed with replicates as measured by CAT assays and normalized to β -galactosidase activity. pCAT-Control (pCATC), which has both the promoter and enhancer of SV40, was used as a positive control, whereas pCAT-Basic (pCATB), which lacks promoter and enhancer, was used as negative control. The relative levels of CAT expression compared with pCAT3 and the wild-type promoter sequence for each construct are illustrated on the right. p3ΔT-CAT represents a mutated promoter in which 5'-TTTATA-3' was changed to 5'-TCTCTA-3', and p3NT-CAT represents a mutated promoter in which the 5'-TTTATA-3' sequence is inverted relative to the canonical 5'-TATAAA-3' sequence.

with TBP with lower affinity, the binding affinity between TBP and the inverted TATA box (K_d 2×10^{-9} M) was comparable with that observed with the consensus TATA box (K_d approx. 2×10^{-9} M) [14].

Disruption of binding of TBP to the inverted TATA element by site-directed mutagenesis

The inverted TATA elements in the rat and human BSP gene promoters were located at the position corresponding to a TATA box in polymerase II-transcribed genes (approx. -30 nt). To investigate whether this inverted TATA element was critical for TBP binding, we created point mutations in the inverted TATA element, either changing 5'-TTTATA-3' into 5'-TCTCTA-3' (p3ΔT-CAT, mutant) to inactivate potential TBP binding, or changing 5'-TTTATA-3' into 5'-TATAAA-3' (p3NT-CAT, 'normal' TATA box) to produce a canonical TATA box. Two site-directed point mutations in the inverted TATA sequence (mutating 5'-TTTATA-3' into 5'-TCTCTA-3') abrogated TBP binding (Figure 3), demonstrating that the inverted TATA element is crucial for the TBP-DNA interaction in the BSP gene. However, conversion of the inverted TATA sequence into a canonical TATA box (5'-TATAAA-3') did not affect TBP binding (K_d 2.2×10^{-9} M) to the template significantly (Figure 3), nor did this change reduce the affinity of the VDRE for the vitamin D₃ receptor (VDR) protein (results not shown).

Requirement of the inverted TATA box for directing downstream expression of the BSP gene

In a previous study we demonstrated basal promoter activity in the rat BSP gene, measured by bacterial CAT reporter gene expression, within a sequence -116 to +60, which encompasses the inverted TATA element and an inverted CCAAT box (shown in Figure 1) [26]. To determine the importance of the inverted TATA sequence in downstream transcription, the chimaeric constructs with the mutated TATA element described above were transiently transfected into the osteoblastic cell line ROS17/2.8 and promoter activity was measured by CAT assays. In accordance with the TBP-DNA-binding data, mutating 5'-TTTATA-3' to 5'-TCTCTA-3' resulted in a marked reduction (by 60–70%) of promoter activity (Figure 4). However, converting the inverted TATA sequence into a canonical TATA box reduced promoter activity only slightly (Figure 4).

Figure 5 *In vitro* transcription of two rat BSP promoter templates

In vitro transcription of two rat BSP promoter templates generated a 171 bp transcript (a) and a 311 bp transcript (b), as indicated by arrows. In both (a) and (b), lane 1, molecular-mass markers; lane 2, mRNA transcribed from wild-type rat BSP promoter; lane 3, mRNA transcribed from mutated rat BSP promoter with a canonical TATA box (changing 5'-TTTATA-3' to 5'-TATAAA-3'); lane 4, mRNA transcribed from rat BSP promoter with point mutations in the inverted TATA element (changing 5'-TTTATA-3' to 5'-TCTCTA-3').

These differences in promoter activity *in vivo* shown in transfection assays were also demonstrated *in vitro* by run-off transcriptions (Figures 5a and 5b). When two DNA templates containing the rat BSP promoter with different run-off transcript sites were incubated with HeLa-cell nuclear extract in the presence of nucleotides and [α -³²P]GTP, radiolabelled run-off transcripts corresponding in size (117 bp and 311 bp) to the products expected for downstream transcription were identified (Figures 5a and 5b). As in the CAT assays (Figure 4), pCAT3 and p3NT-CAT showed similar transcription activity in this system whereas no transcription activity was observed with the p3ΔT-CAT mutant.

TBP is required for initiation of the BSP gene promoter in transcription

We also performed *in vitro* transcription experiments to determine whether TBP is an absolute requirement for the initiation of transcription driven by the BSP promoter. We first inactivated

Figure 6 Reconstitution of *in vitro* transcription directed by rat BSP promoter with recombinant human TBP

Lane 1, molecular-mass markers; lane 2, 171 bp transcript generated by nuclear extract from HeLa cells (indicated by arrow); lane 3, *in vitro* transcription of the rat BSP promoter using heat-treated HeLa-cell nuclear extract; lane 4, *in vitro* transcription of the rat BSP promoter using heat-treated HeLa-cell nuclear extract plus 20 ng of BSA; lanes 5–7, *in vitro* transcription of the rat BSP promoter using heat-treated HeLa-cell nuclear extract plus 10, 20 and 40 ng of recombinant human TBP respectively.

the TBP by heating the HeLa-cell nuclear extract at 45 °C for 15 min [34] and showed that the nuclear extract containing heat-inactivated TBP was not able to initiate transcription when incubated with wild-type rat BSP promoter–CAT construct (pCAT3) (Figure 6). However, we subsequently showed that *in vitro* transcription could be rescued by adding recombinant human TBP to the heat-treated nuclear extract (Figure 6).

DISCUSSION

Although several variants of the classical TATA sequence have been observed, most of the sequences are asymmetrical [7,14,15], implying that the orientation of the TATA box may be important for directing downstream transcription, as discussed in a number of recent reviews [1, 15–17,22,38]. The nucleotide sequence in the TATA box of the BSP gene is inverted relative to the consensus TATA box, yet it is still functional and crucial in driving downstream transcription of the BSP gene. By characterizing the naturally occurring inverted TATA box in the BSP gene, we provide conclusive evidence that the direction of transcription is not dependent on the orientation of the TATA box in eukaryotic genes.

An apparently inverted TATA-box sequence (5'-TTTGTA-3') has been previously reported in the adenovirus IVa2 promoter [39,40]. However, in the IVa2 gene, the TATA box is located 15 nucleotides downstream of the transcription start site. Although synthetic promoters with both canonical and inverted TATA sequences have been shown to drive transcription from similar initiation sites and in the same direction when used in an *in vitro* transcription system, transcription activities from the artificially inverted TATA sequence are greatly reduced [20,41]. In the BSP gene promoter, transcription activities were comparable between the wild-type inverted TATA sequences and the mutated canonical TATA box both *in vitro* and *in vivo*. As in all other RNA polymerase II-transcribed genes, either with or

without a TATA box, the activity of the BSP promoter containing a perfect inverted TATA box required the participation of TBP. Thus it is apparent that TATA boxes, like enhancers and silencers, are operational in either orientation.

The concept that the orientation of the TATA box directs the polarity of transcription was based on the observation that most TATA boxes discovered so far are asymmetric in nucleotide sequences [1,15–17,22,38]. In addition, inversion of a TATA box artificially results in loss of downstream transcription activity [18,19]. The elucidation of the tertiary structure of TBP from the plant *Arabidopsis thaliana* [42] has revealed that TBP resembles a molecular saddle which could sit astride the duplex DNA, protecting approx. 10 bp of DNA sequence. Although the two functional domains of TBP are topologically symmetrical, the amino acids forming the surfaces that interact with the DNA are different [42], as predicted by previous mutagenesis studies of TBP [15,43,44]. Consequently, it is likely that TBP interacts with the TATA element in a directional manner [15,16,42]. More recently, the three-dimensional structure of plant and yeast TBPs complexed with the typical TATA boxes were determined by X-ray crystallography [38,45]. In both cases, the TBPs interacted with the TATA boxes with the same polarity, and the TBPs bound to the minor groove of DNA as predicted by earlier studies [36,37]. On the basis of these data, Kim et al. [38] and Klug [22] hypothesized that 'the preferred orientation of TBP/TATA binding may be derived from asymmetry in the deformability of the TATA element'. However, the possibility that the same polarity of TBP/TATA binding in both cases of co-crystallization might be achieved as the result of a random process still cannot be excluded. Our present study indicates that the questions of TBP/TATA binding polarity and the transcription directionality are more complicated than previously thought. It is of interest to note that the initiator element has also been suggested to be a determinant for the direction of transcription [20]. However, a consensus sequence for this element is not found in either the rat or the human BSP promoter.

The three-dimensional structure of the TBP from *A. thaliana* complexed with a 14 bp oligonucleotide bearing the AdMLP TATA box has recently been refined at 0.19 nm (1.9 Å) resolution by Kim and Burley [46]. Model building based on their newly refined co-crystal structures predicted that substitution of a G-C base pair at position 2 reduces TATA box efficiency to 1% whereas a T-A substitution at this position would be tolerated. Inversion of the AdMLP TATA box generates an inverted TATA box with T-A to A-T substitutions at positions 2, 4 and 6, and inversion of the yeast *CYC1* TATA box generates A-T to T-A substitutions at positions 2 and 7. Model-building studies suggested that these inverted TATA boxes do not generate obviously unfavourable interactions with TBP. These studies may therefore explain our observations that the inverted TATA box of the BSP gene maintains a comparable activity with that of a canonical TATA box. Of note, the inverted TATA box of the BSP gene differs from the above-mentioned inverted TATA box in that the T-A substitutions are located at positions 2 and 5. Although the binding polarity could be due to asymmetry in the deformability of the two domains of TBP and the two halves of the recognition site of the TATA box [38,46], how TBP actually binds to the inverted TATA box and how the direction of transcription is achieved is yet to be resolved. Thus it would be of interest to co-crystallize TBP with the functional inverted TATA box of the BSP gene. The inverted TATA box in the BSP gene promoter therefore can provide an excellent model to facilitate studies of the polarity of the TFIID–DNA complex and its relation to downstream transcription.

The ability of an asymmetrical TBP molecule to bind to an

Figure 7 Model depicting possible interactions between the TBP and the inverted TATA box

In (a) the flanking DNA sequence of the inverted TATA box (5'-TTTATA-3') or the auxiliary factors (e.g. TAFs) determines TBP binding to the upper strand, and polymerase II is thus orientated for downstream transcription. In (b) the TBP binds to the 5'-TATAAA-3' sequence on the lower strand, but TAFs or other auxiliary factors may orientate polymerase II for downstream transcription.

inverted TATA box and to form a functional PIC has important implications in elucidating the binding interactions between the promoter, TBP and PIC components. Two possible models are presented (Figure 7) to explain the functional implications with respect to the binding of TBP to an inverted TATA box. First, TBP binding to the 5'-TTTATA-3' sequence on the upper strand of the inverted TATA box could be modulated by the flanking sequences of the TATA box or by the auxiliary factors (e.g. TAFs) as depicted in Figure 7(a), and the downstream transcription is thus determined. In this model, the binding of TBP and its orientation would be the same as in the consensus TATA box and variant TATA-box sequences, and perhaps also in TATA-less promoters. As an example, nucleotides flanking a conserved TATA core are capable of dictating the DNA binding of murine homeodomain proteins [47]. Secondly, TBP could bind to the 5'-TATAAA-3' sequence on the lower strand of the inverted TATA box. In this model the positioning of the RNA polymerase II for downstream transcription might be determined by the auxiliary factors such as the TAFs (Figure 7b), in which case protein-DNA interactions other than those involved with TBP binding would be expected. In this regard there is evidence indicating that the direction of transcription might depend on the relative positions of the TATA box and other upstream elements such as the CCAAT box [21]. Notably, an inverted CCAAT box that is required for basal promoter activity is retained within the highly conserved immediate promoter region of the rat and human BSP [26,27].

The functional significance of an inverted TATA box in the BSP gene promoter still remains to be investigated. Recently, we have identified a VDRE that overlaps this inverted TATA box and mediates the repression effect of 1,25-dihydroxyvitamin D₃ on the rat BSP gene [48]. As demonstrated by gel mobility-shift analysis, both endogenous and recombinant VDRs bind to this putative rat BSP VDRE in a concentration-dependent manner with a strong preference for dimer formation (R. H. Kim, J. J. Li, Q. Zhang, Y. Ogata, M. Yamauchi, L. P. Freedman and J. Sodek, unpublished work). The fact that both TBP and VDR can bind to the same DNA segment indicates a novel mechanism for gene suppression by vitamin D₃ whereby VDR could compete with TBP for occupancy of the TATA/VDRE and prevent the

assembly of the PIC. However, that the binding affinity of VDR is not reduced after the mutation of this sequence to produce a normal TATA box indicates that the inverted TATA sequence is not required to accommodate the VDRE.

Variant TATA boxes have been suggested to function as a selective element to determine tissue-specific gene expression by interaction with distinctive TFIID complexes [3,8,49]. As a variant form of the TATA box, it is possible that the inverted TATA box may play an important role in controlling the spatial and temporal expression of the BSP gene during bone formation and development. It is conceivable that the variant sequences of the TATA boxes might be important for the recognition of certain transcription factors that compete with TBP for binding in this region and thereby regulate gene expression. Moreover, as the binding of TBP to the TATA box induces the DNA to bend [38,45], variant nucleotide(s) in the TATA boxes might be able to dictate the degree of DNA bending, and as such provide a means of fine tuning by stimulation or repression of transcription via TBP-specific *trans*-factor interactions. As it is recognized that TBP is able to interact with a number of transcription factors [8,16], it is also conceivable that the specificity of such interactions might depend on the specific changes of TBP-DNA conformation on the binding of TBP to various types of TATA boxes.

We thank Dr. J. Greenblatt for advice and providing recombinant yeast TBP for our initial studies, and Dr. G. A. Rodan for the ROS17/2.8 cell line. We are grateful to Ms. K.-L. Yao and Dr. H. K. Moe for technical assistance. We also appreciate the stimulating discussions of these studies with Dr. S. K. Burley, Dr. V. Giguère, Dr. J. L. Wrana, Dr. Y. Ogata and Dr. Q. Zhang. This investigation was supported by a Medical Research Council (Canada) Group grant. J. J. L. is the recipient of an I'Anson Fellowship.

REFERENCES

- Greenblatt, J. (1991) *Cell* **66**, 1067-1070
- Sawadogo, M. and Sentenac, A. (1990) *Annu. Rev. Biochem.* **59**, 711-754
- Pugh, B. F. and Tjian, R. (1992) *J. Biol. Chem.* **267**, 679-682
- Wiley, S. R., Kraus, R. J. and Mertz, J. E. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 5814-5818
- Dynlacht, B. D., Hoey, T. and Tjian, R. (1991) *Cell* **66**, 563-576
- Takada, R., Nakatani, Y., Hoffmann, A. et al. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 11809-11813
- Breathnach, R. and Chambon, P. (1981) *Annu. Rev. Biochem.* **50**, 349-383
- Brou, C., Chaudhary, S., Davidson, I. et al. (1993) *EMBO J.* **12**, 489-499
- Buratowski, S., Hahn, S., Guarente, L. and Sharp, P. A. (1989) *Cell* **56**, 549-561
- Maniatis, T., Goodbourn, S. and Fischer, J. A. (1987) *Science* **236**, 1237-1245
- Mitchell, P. J. and Tjian, R. (1989) *Science* **245**, 371-378
- Ptashne, M. and Gann, A. A. F. (1990) *Nature (London)* **346**, 329-331
- Stringer, K. F., Ingles, C. J. and Greenblatt, J. (1990) *Nature (London)* **345**, 783-786
- Hahn, S., Buratowski, S., Sharp, P. A. and Guarente, L. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 5718-5722
- Reddy, P. and Hahn, S. (1991) *Cell* **65**, 349-357
- Greenblatt, J. (1992) *Nature (London)* **360**, 16-17
- Phillips, S. E. V. (1993) *Curr. Biol.* **3**, 112-114
- Nagawa, F. and Fink, G. R. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 8557-8561
- Ruden, D. M., Ma, J. and Ptashne, M. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4262-4266
- O'Shea-Greenfield, A. and Smale, S. T. (1992) *J. Biol. Chem.* **267**, 1391-1402
- Xu, L., Thali, M. and Schaffner, W. (1992) *Nucleic Acids Res.* **19**, 6699-6704
- Klug, A. (1993) *Nature (London)* **365**, 486-487
- Sodek, J., Chen, J., Kasugai, S. et al. (1992) in *Chemistry and Biology of Mineralized Tissues* (Slavkin, H. and Price, P., eds.), pp. 297-306, Elsevier Science Publishers, Amsterdam
- Sodek, J., Chen, J., Kasugai, S. et al. (1992) in *The Biological Mechanisms of Tooth Movement and Craniofacial Adaptation* (Davidovitch, Z., ed.), pp. 127-136, EBSCO Media, Birmingham, AL
- Hunter, G. K. and Goldberg, H. A. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 8562-8565
- Li, J. J. and Sodek, J. (1993) *Biochem. J.* **289**, 625-629
- Kim, R. H., Shapiro, H. S., Li, J. J., Wrana, J. L. and Sodek, J. (1994) *Matrix Biol.* **14**, 31-34

-
- 28 Kerr, J. M., Fisher, L. W., Termine, J. D. et al. (1993) *Genomics* **17**, 408–415
- 29 Horikoshi, M., Wang, C. K., Fujii, H., Cromlish, J. A., Weil, P. A. and Roeder, R. G. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 4843–4847
- 30 Ausubel, F. M., Brent, R., Kingston, R. E. et al. (1992) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York
- 31 Landt, O., Grunert, H.-P. and Hahn, U. (1990) *Gene* **96**, 125–128
- 32 Farnham, P. and Schimke, R. T. (1986) *Mol. Cell Biol.* **6**, 2392–2401
- 33 Innis, J. W., Moore, D. J., Kash, S. F., Ramamurthy, V., Sawadogo, M. and Kellems, R. E. (1991) *J. Biol. Chem.* **266**, 21765–21772
- 34 Nakajima, N., Horikoshi, M. and Roeder, R. G. (1988) *Mol. Cell. Biol.* **8**, 4028–4040
- 35 Hoey, T., Dynlacht, B. D., Peterson, M. G., Pugh, B. F. and Tjian, R. (1990) *Cell* **61**, 1179–1186
- 36 Lee, D. K., Horikoshi, M. and Roeder, R. G. (1991) *Cell* **67**, 1241–1250
- 37 Starr, D. B. and Hawley, D. K. (1991) *Cell* **67**, 1231–1240
- 38 Kim, Y., Geiger, J. H., Hahn, S. and Sigler, P. B. (1993) *Nature (London)* **365**, 512–520
- 39 Carcamo, J., Maldonado, E., Cortes, P. et al. (1990) *Genes Dev.* **4**, 1611–1622
- 40 Kasai, Y., Chen, H. and Flint, S. J. (1992) *Mol. Cell. Biol.* **12**, 2884–2897
- 41 Carcamo, J., Buckbinder, L. and Reinberg, D. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 8052–8056
- 42 Nikolov, D. B., Hu, S. H., Lin, J. et al. (1992) *Nature (London)* **360**, 40–46
- 43 Strubin, M. and Struhl, K. (1992) *Cell* **68**, 721–730
- 44 Yamamoto, T., Horikoshi, M., Wang, J., Hasegawa, S., Weil, P. A. and Roeder, R. G. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2844–2848
- 45 Kim, J. L., Nikolov, D. B. and Burley, S. K. (1993) *Nature (London)* **365**, 520–527
- 46 Kim, J. L. and Burley, S. K. (1994) *Nature, Struct. Biol.* **1**, 638–653
- 47 Catron, K. M., Iler, N. and Abate, C. (1993) *Mol. Cell. Biol.* **13**, 2354–2365
- 48 Sodek, J., Li, J. J., Kim, R. H. et al. (1994) in *The Biological Mechanisms of Tooth Eruption, Resorption and Replacement by Implants* (Davidovitch, Z., ed.), pp. 215–226, EBSCO Media, Birmingham, AL
- 49 Wefald, F. C., Devlin, B. H. and Williams, R. S. (1990) *Nature (London)* **344**, 260–262
-