Stereochromical Analysis of the Functional Significance of the Conserved Inverted CCAAT and TATA Elements in the Rat Bone Sialoprotein Gene Promoter*

Received for publication, July 29, 2005, and in revised form, January 5, 2006. Published, JBC Papers in Press, February 22, 2006. DOI 10.1074/jbc.M508364200

Ming Su‡, Daniel Lee*, Bernhard Ganss†, and Jaro Sodek‡§

From the ‡Canadian Institutes of Health Research Group in Matrix Dynamics, Faculty of Dentistry and §Department of Biochemistry, Faculty of Medicine, University of Toronto, Toronto, Ontario M5S 3E2, Canada

Basal transcription of the bone sialoprotein gene is mediated by highly conserved inverted CCAAT (ICE; ATTG) and TATA elements (TTTATA) separated by precisely 21 nucleotides. Here we studied the importance of the relative position and orientation of the CCAAT and TATA elements in the proximal promoter by measuring the transcriptional activity of a series of mutated reporter constructs in transient transfection assays. Whereas inverting the TTTATA (wild type) to a TATAAA (consensus TATA) sequence increased transcription slightly, transcription was reduced when the flanking dinucleotides were also inverted. In contrast, reversing the ATTG (wild type; ICE) to a CCAAT (RICE) sequence caused a marked reduction in transcription, whereas both transcription and NF-Y binding were progressively increased with the simultaneous inversion of flanking nucleotides (f-RICE-f). Reducing the distance between the ICE and TATA elements produced cyclical changes in transcriptional activity that correlated with progressive alterations in the relative positions of the CCAAT and TATA elements on the face of the DNA helix. Minimal transcription was observed after 5 nucleotides were deleted (equivalent to approximately one half turn of the helix), whereas transcription was fully restored after deleting 10 nucleotides (approximately one full turn of the DNA helix), transcriptional activity being progressively lost with deletions beyond 10 nucleotides. In comparison, when deletions were made with the ICE in the reversed (f-RICE-f) orientation transcriptional activity was progressively lost with no recovery. These results show that, although transcription can still occur when the CCAAT box is reversed and/or displaced relative to the TATA box, the activity is dependent upon the flexibility of the intervening DNA helix needed to align the NF-Y complex on the CCAAT box with preinitiation complex proteins that bind to the TATA box. Thus, the precise location and orientation of the CCAAT element is necessary for optimizing basal transcription of the bone sialoprotein gene.

Regulation of class II nuclear gene promoters involves interaction between transcriptional activators and coactivators and accessory proteins with general transcription factors that form the preinitiation complex (PIC)1 (1, 2). Many polymerase II promoters contain a TATA box (3) through which TFIIID binds and directs the assembly of other tran-

9882 JOURNAL OF BIOLOGICAL CHEMISTRY
This is an Open Access article under the CC BY license.

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

1 To whom correspondence should be addressed: CIHR Group in Matrix Dynamics, Faculty of Dentistry, 234 FitzGerald Blvd., 150 College St., University of Toronto, Toronto, ON M5S 3E2, Canada. Tel: 416-978-6624; Fax: 416-978-3956; E-mail: msu@utoronto.ca.

‡ The abbreviations used are: PIC, preinitiation complex; BSP, bone sialoprotein; ICE, inverted CCAAT element; RICE, reversed ICE; TBP, TATA-binding protein; nts, nucleotides; EMSA, electrophoretic mobility shift assay; ChiP, chromatin immunoprecipitation; OPN, osteopontin; DSP, dentin sialophosphoprotein; WT, wild type.

© 2006 by The American Society for Biochemistry and Molecular Biology, Inc. Printed in the U.S.A.
**MATERIALS AND METHODS**

**Cell Culture**—Rat osteosarcoma cells (ROS 17/2.8) (provided by Dr. Gideon Rodan, Merck-Frost, Philadelphia, PA) and RBMC-D8 cells were grown in α-minimal essential medium containing 10% fetal bovine serum and antibiotics (penicillin G, 100 μg/ml; gentamicin sulfate, 50 μg/ml; and fungizone, 25 μg/ml).

Preparation of Rat BSP Promoter Constructs—All of the constructs used in this study are based on ~600-BSP Luc, in which ~60 to ~600 of the rat BSP promoter sequence was blunt end-ligated into Smal site of pGL3-Basic (Promega, Madison, WI). The following constructs were generated using the QuickChange site-directed mutagenesis kit (Stratagene, La Jolla, CA): constructs with a single point-mutated CCAAT box, rTTGG; with a reversed CCAAT box, RICE; with the TATA element inverted, tataaa; with both the wild type ICE and TATA reversed, RICE + tataaa; and with 2 nts on each side of the TATA inverted (from GG to CC) with the TATA box, RvT-RITATA; and with both the ICE and the flanking nts of TATA reversed, RICE + RITATA. To prepare 1) a series of promoter constructs with deletions between the ICE (and mutated ICE) and TATA boxes; 2) a series of ICE with flanking sequences inverted; 3) the reversed ICE with 5 flanking nucleotides, (f5-RICE-f5); 4) the ICE in a 3′ → 5′ orientation with five flanking nucleotides, (f5-FlipRICE-f5); and 5) the RICE in a 3′ → 5′ orientation with five flanking nucleotides (f5-FlipRICE-f5), −60 BSP Luc was first digested with Nhel/Xhol. The 130-bp Nhel/Xhol fragment was then further digested with MboII to obtain the 84-bp MboII/XhoI “adapter” fragment. Oligonucleotides with 5′-Nhel sites and 3′-MboII sites and with those required mutations were ordered from Invitrogen Canada Inc. (Burlington, Canada). The annealed complementary oligonucleotides, the 84-bp adapter, and the Nhel/Xhol-digested and dephosphorylated −60 BSP Luc were then ligated at +16 °C overnight. All of the constructs were confirmed by sequencing.

**Transcription Assays**—ROS 17/2.8 and RBMC-D8 cells were plated on 24-well cell culture plates (0.4 × 10^5/well) 24 h prior to transfection. The cells (60% confluent) were then transfected using Lipofectamine 2000 (Invitrogen). The double-stranded oligonucleotides were end-labeled with [γ-^32P]ATP (PerkinElmer Life Sciences) using T4 polynucleotide kinase (Invitrogen). 2.5 μg of nuclear extracts, or 1 μg of recombinant NF-Y were incubated with 50,000 dpm of labeled oligonucleotides at room temperature (21 °C). The nuclear extracts were incubated with or without antibody to NF-Y (23) at room temperature for 10 min, and then the probe was added and incubated for an additional 20 min. The sequences of the probes used are (only the sense strands are shown) as follows: s-CCAAAT (ICE), CCGTGACCGTATGGGCT-GCTGAGAG; RICE, CCGTGACCGTGCCcaataCCTGAGAG; f5-RICE-f5, CCGTGACcagcccaatacagcAGG; f5-RICE-f5, CCGTccagcagcccaatacagcAGG; f7-RICE-f7, CCGTtcagcagcccaatacagcAGG; f8-RICE-f8, CCGrctagcagcccaatacagcAGG; f9-RICE-f9, CcctagcagcagcccaatacagcAGG; and RITATA, CcctagcagcagcccaatacagcAGG. The bound DNA was PCR-amplified using promoter-specific primers, resolved by agarose gel electrophoresis, and then visualized by UV fluorescence.

**Chromatin Immunoprecipitation Assays (ChIP)**—The ChIP procedure was adapted from Mao et al. (25). HeLa cells were plated on 15-cm plates 24 h prior to transfection. For each pull-down, 1 × 10^6 cells were used. HeLa cells grown to ~60% confluence were transfected with total DNA of 18.6 μg/plate (~600 Luc, rTTGG, RICE, or D10), using 93 μl/15-cm plate Lipofectamine 2000 (Invitrogen). The cells were incubated for 44 h and fixed in 1% (v/v) formaldehyde. The extracted chromatin was sonicated (7 × 10 s at 13 W to fragment DNA to an average size of 600 bp) and then immunoprecipitated using antibodies that recognize NF-YA (CBE-B (H-209); Santa Cruz Biotechnology, Santa Cruz, CA) or normal rabbit serum as a negative control. The chromatin immunoprecipitate was de-cross-linked with 0.3 M of NaCl. The bound DNA was PCR-amplified using promoter-specific primers, resolved by agarose gel electrophoresis, and then visualized by UV fluorescence.

**RESULTS**

**Conservation of the Mammalian BSP Promoters**—Comparison of the proximal promoter sequences of the human, rat, and mouse BSP genes has shown a highly conserved ICE and an inverted TATA box, which are separated by 21 nucleotides (18). In addition, there is high sequence identity in the promoter sequences, particularly in the regions flanking the pentanucleotide CCAAT element. Notably, other SIBLING (small integrin-binding ligand N-linked glycoprotein) family members, rodent osteopontin (OPN) and dentin sialophosphoprotein (DSP) genes also include an inverted CCAAT box upstream from an atypical TATA box (Fig. 1). Whereas the TATA box in BSP is perfectly inverted relative to the consensus TATAAA sequence, in the TATA boxes of the OPN and DSP promoters the penultimate thymidine is replaced by an adenine. In the OPN gene the distance between the ICE and TATA elements is also 21 nts, whereas these elements are separated by 22 nts in the DSP gene. To study the importance of the orientation and relative positions of the CCAAT and TATA elements, we analyzed the effects of a series of mutations on the transcriptional activity of a short promoter construct pBSP +60 to −60 ligated to a luciferase report gene. The strategy employed in these studies is depicted in Fig. 2.
Stereochemical Relationship between CCAAT and TATA Elements

Orientation of the CCAAT and TATA Boxes—To determine the importance of the orientation of the CCAAT and TATA boxes in the basal transcription of the BSP gene, mutations were introduced to reverse the orientation of these elements both separately and together. When the pentanucleotide ICE box alone was reversed (RICE) to a 5'→3' CCAAT sequence, transcription measured in ROS 17/2.8 osteosarcoma cells was markedly reduced but not to the extent observed with a single nucleotide mutation in the ICE box, whereas reversing the TATA resulted in a modest but reproducible increase in transcription (Fig. 2), as found previously (20). When both elements were reversed, transcription was reduced but not to the same extent as reversing the ICE box alone, probably because of the more favorable effect of the change in the TATA box orientation. As anticipated, no transcriptional activity was obtained when either the CCAAT or ICE sequence was placed in a 3'→5' orientation.

Flanking Sequences and Transcription—To determine the importance of the flanking sequences, two nucleotides on each side of the TTTATA element were reversed with the inverted TATA sequence (Fig. 3). This change in flanking nucleotides reduced transcription with the ICE box left intact and more extensively when the orientation of the ICE was reversed (i.e. ICE → RICE). In comparison, when the flanking sequence was reversed with the ICE (f-ICE-f), transcription was progressively increased with an increase of flanking nucleotides until 5 nucleotides on each side had been reversed. With 5 flanking nucleotides reversed with the ICE, transcription approached the rate for the control WT promoter. However, reversing additional flanking nucleotides, beyond 5 nucleotides, resulted in decreased transcription (Fig. 3). Notably, the three nucleotides on the 5' side and on the 3' side in the reversed sequence conform to the consensus sequence determined for NF-Y binding (Fig. 1). When cells were cotransfected with an expression vector for the NF-YA subunit, luciferase activities were increased proportionately for each construct, indicating an enhancement of the NF-Y-mediated transcription.

Thus, although the orientation of the TATA element is not critical for TBP binding, the nature of the flanking nucleotides appear to be important. In contrast to the TATA box, the flanking sequence must be reversed with the ICE sequence to obtain a functional CCAAT box.

Flanking Sequences and EMSA—To determine whether the reversed flanking sequences were important for retaining NF-Y interactions, transcription factor binding was assessed by EMSA using nuclear extracts from ROS 17/2.8 cells (Fig. 4). With control oligonucleotides encompassing the ICE box and its normal flanking sequences (WT), gel shifts identified a strong band that was efficiently supershifted with antibodies to NF-YB and a minor more quickly migrating band that was unaffected by NF-Y antibodies (26). With the RICE sequence, a new major band was seen together with a diffuse faster migrating band that migrated differently from the original minor bands. However, with only two 5'- and 3'-flanking nucleotides reversed with the RICE, the EMSA pattern showed the original minor bands and the major band shifted with NF-Y antibodies, albeit in a lower amount than seen with...
the WT probe. Further increases in reversed flanking sequence produced progressive increases in NF-Y binding, with maximal binding seen with 10 flanking nucleotides reversed. The progressive increase in NF-Y binding was also evident when shifts were performed with recombinant NF-Y, confirming that the major band in the nuclear shifts was NF-Y and that maximal binding of NF-Y occurred with 10 flanking nucleotides. Although the increases in NF-Y binding corresponded to increases in transcription with <5 nucleotides reversed, thereafter tran-
Increased transcription in all of the deleted constructs (Fig. 5) shows that NF-Y-mediated transcription was maintained following the deletion of 5′ nucleotides. However, transcription was progressively and completely lost. Notably, the decrease after 12 deletions was fully restored and was frequently above that observed with the WT promoter. Beyond 10 deletions, there was a further decrease in transcriptional activity. Thus, after 12 deletions occurred despite the retention of 5′ flanking nucleotides on each side of the ICE box, suggesting that the decrease was not due to compromised transcription factor recognition. To demonstrate that NF-Y-mediated transcription was maintained following the deletion of nucleotides, ROS 17/2.8 cells were cotransfected with NF-YA, which increased transcription in all of the deleted constructs (Fig. 5B). The study was also replicated in a normal rat bone marrow cell line, RBMC-D8, in which cotransfection with a NF-Y-DN vector suppressed the transcriptional activity consistent with NF-YA-mediated transcription (Fig. 5C). Notably, the degree of stimulation by NF-YA and suppression by NF-Y-DN was variable, which may reflect the effect of positional differences on relative expression levels.

To confirm that NF-Y binds to the CCAAT box in the undeleted and deleted constructs in vivo, chromatin immunoprecipitation analyses were conducted in HeLa cells, which express high levels of NF-Y. Following transfection with the WT construct or the 10-nucleotide-deleted construct (D10), promoter complexes were immunoprecipitated with antibodies to NF-Y (αCBF-B) and a promoter sequence encompassing the CCAAT region amplified by PCR. A band corresponding to the input DNA was strongly amplified for the WT and D10 constructs (Fig. 6), whereas only a weak amplification was observed with transfected promoter constructs in which the CCAAT box was either reversed (RICE) or included a single base mutation (tTTGG). These studies show that NF-Y binding in vivo is retained in the deletion construct.

Separation of the ICE and TATA Boxes—To determine the relationship between the orientation of the CCAAT box and its distance from the TATA box, deletions were introduced between the TATA and 5′ RICE-f5, in which transcriptional activity is essentially optimal for the reversed CCAAT element. With the ICE sequence reversed, transcription was also reduced progressively with nucleotide deletions (Fig. 7). However, contrasting the results with the WT promoter, the transcriptional activity was further reduced beyond 5 deletions with minimal transcription observed after 10 deletions. To complete these analyses we also examined transcriptional activity when the ICE box sequences, together with 5′ flanking nucleotides on each side were reversed (flipped) in direction from their 5′ → 3′ orientation to a 3′ → 5′ orientation (Fig. 2). In both cases, transcriptional activity was lost, and no effect was observed when intervening nucleotides were deleted (Fig. 7), indicating that the 5′ → 3′ direction of the CCAAT box (whether in the ICE or 5′ RICE-f5 orientation) is critical for transcription factor binding.

The results obtained in this series of experiments could be explained according to the stereochemical relationship between the orientation and position of the ICE and TATA elements, as shown in Fig. 8. Although the orientation of the CCAAT-binding transcription factor is retained, its position relative to the TATA box is changed as deletions are introduced; each deletion not only shortening the distance from the TATA box, but also changing its relative position on the face of the DNA helix. Thus, after five deletions the ICE box has rotated ~180°, and its position is maximally removed from alignment with the TATA, and minimal transcription activity is seen. Although the rate of transcription is compromised, presumably because of the energy required to twist the DNA, there would appear to be sufficient distance to allow the intervening DNA to bend and thereby allow interaction between the CCAAT-binding transcription factor complex and PIC proteins. When the ICE sequence is reversed with its flanking sequence (f5-RICE-f5), the DNA is distorted to allow the RICE-binding transcription factor to “flip” orientation and interact with the PIC complex. This appears to be possible because transcriptional activity is similar to the WT constructs and indeed loss of transcription with five nucleotide deletions is similar to the reduction with the WT promoter construct. However, with further deletions there is insufficient intervening DNA to allow the flip to occur, so recovery of transcription, when the RICE and TATA are realigned at 10 deletions, is not observed.

**DISCUSSION**

These studies have investigated the importance of the stereochemical relationship between the ICE and inverted TATA elements that are required for basal transcription and whose position and orientation are conserved in a large number of eukaryotic genes, including BSP and SIBLING family members. Although reversing the inverted TATA into a classical TATAA element resulted in a modest increase in transcription, a similar reversal of the ICE box sequence (RICE) caused a loss of transcriptional activity. However, transcriptional activity could be recovered by reversing the flanking nucleotides with the ICE. By deleting nucleotides to reduce the separation of the ICE and TATA boxes, transcription was progressively reduced and subsequently recovered, reaching maximal transcription after the deletion of 10 nts, corresponding to a single turn of the DNA helix. These results suggest that the change in relative positions of the ICE and TATA boxes on the face of the helix compromises interactions between transcription factors that associate with these elements, but that optimal interactions can be achieved with 11 nucleotides separating the elements. In contrast, when the ICE box was reversed together with its flanking sequence (f5-RICE-
f5) there was no recovery of transcriptional activity as nucleotides were deleted. Collectively, these results indicate that a minimal length of DNA is needed to provide the flexibility for the TATA and RICE boxes to realign and orientate into a favorable position for transcription factor interactions.

An inverted TATA box was originally described for the AdIVa2 promoter, which is present downstream of the transcription initiation site (27). However, the inverted TATA sequence in BSP, which overlaps a vitamin D3 response element (28), was the first identified in eukaryotic genes (20). Although an inverted TATA box in synthetic promoters had previously been shown to direct transcription in vitro, the transcription efficiency was markedly reduced (29, 30). However, transcriptional activity in the BSP gene is only slightly lower (∼10%) than when the sequence is converted into a "consensus" TATAAA sequence (20) (Fig. 2). Although the TBP could bind to the inverted TATA box in the opposite orientation on the lower strand, the structural basis of the apparent directionality of TBP binding remains a subject of much discussion and speculation (6, 31). Studies of the interaction between TBP
Stereochemical Relationship between CCAAT and TATA Elements

and the TATA box show that A → T and T → A substitutions at positions 2, 4, and 6 in the TATAAA sequence can be tolerated, assuming that the inverted TATA box can undergo the same conformational change as observed with the adenovirus major late promoter and yeast CYC1 TATA boxes (3) used in the crystallographic studies of TBP-TATA box interactions (32, 33). However, in crystallographic analyses TBP binds in the “normal” orientation, and it has been suggested that the slightly different curvatures of the two halves of the TBP β-sheet determine directionality (34). Although Juo et al. (35) have suggested that proline 285 (human TBP numbering) determines the direction of TBP binding through its minor groove interaction with the first T-A base pair, this proposal is not supported by mutational analysis (31). Moreover, in contrast to the crystallographic studies, TBP in solution binds the TATA box in both orientations with only a slight preference for the consensus orientation (36), and there are no unfavorable interactions when TBP binds to the TATA box in the reverse orientation (37).

Notably, the deformation of the TATA element by TBP (32, 33, 38) creates a structurally invariant nucleoprotein complex that serves as the receptor for TFIIB and TFIID (39), which increase the preference for normal orientation of the TBP (40). Thus, although the polarity of transcription is independent of TATA box orientation (8), TBP probably binds in the normal orientation to the inverted TATA box in BSP because a TA-rich sequence with Ts predominant on the 5’ end appears to provide the necessary binding and orientation. When the flanking nucleotides are inverted a C is placed at position 7. Although crystallographic studies indicate that this substitution can be accommodated through the formation of a C:G Hoogsteen base pair (3), transcription is reduced (Fig. 4).

That the sequence of the TATA box can function well in either orientation or with various T → A mutations is consistent with the ability of TBP to recognize variants of nucleotides that form the minor groove of the TATA element, which is believed to be important in preventing potentially lethal mutations without perturbing subsequent steps in PIC assembly (35). However, this conflicts with the frequent conservation of gene specific TATA box sequences, such as the inverted sequence in the ICE/ICE-RICE-f5 elements relative to the PIC as deletions are made is shown in cross-section (C). The effect of reducing the distance between the ICE and TATA on transcription is depicted with bending of the DNA allowing interaction between the transcription factors associated with the two elements (A). As shown in B, for the f5-RICE-f5 to interact it must be flipped as well as rotated as deletions are made. As a result, in contrast to the ICE, transcriptional activity is not recovered after the deletion of 10 nucleotides (approximately one turn of the helix). The change in the orientation of the ICE/ICE-RICE-f5 elements relative to the PIC as deletions are made is shown in cross-section (C).

That immediate upstream activators may be involved in the directionality of TBP binding and transcription is indicated from the topographical analysis of a synthetic promoter in which an ATTTGCACT octamer was found to direct transcriptional polarity through a downstream TATA complex (8). More recently, the Gal4-VP16 and Gal4AH activators were shown to enhance the orientation and axial specificity of the TBP/TATA complex (40). Because the CCAAT-binding protein NF-Y has been shown to be essential for the recruitment of general transcription factors and RNAPII to the core promoter of the γ-globulin (42), OPN and ODF/RANKL genes, as well as to the TATA-less E2F1 gene promoter (43), it is conceivable that directionality of transcription could be determined by the proximal CCAAT element that is present in ~30% of eukaryotic genes (12). Accordingly, the orientation of the TATA sequence would be irrelevant for basal transcription.

FIGURE 6. Chromatin immunoprecipitation assays showing NF-Y binding to the promoter constructs in vivo. HeLa cells were transfected with BSP promoter constructs WT, tTGG, RICE, or D10. The cells were harvested 44 h after transfection, and the chromatin was extracted and sonicated. The chromatin fragments were immunoprecipitated with αCBF-B (NF-YA) or normal rabbit serum as negative control and subjected to PCR amplification using primers encompassing the BSP promoter and luciferase gene. Gel electrophoresis of the PCR products is shown here. The input shown is a 1:300 dilution.

FIGURE 7. Transcription assays of BSP promoter constructs with deletions of 5 or 10 nt from f5-RICE-f5 and f5-FlipC-f5. Constructs with 5- or 10-nt deletions from f5-RICE-f5 or f5-FlipC-f5 were cloned and used in transient transfections of ROS17/2.8 and the TATA box show that A → T and T → A substitutions at positions 2, 4, and 6 in the TATAAA sequence can be tolerated, assuming that the inverted TATA box can undergo the same conformational change as observed with the adenovirus major late promoter and yeast CYC1 TATA boxes (3) used in the crystallographic studies of TBP-TATA box interactions (32, 33). However, in crystallographic analyses TBP binds in the “normal” orientation, and it has been suggested that the slightly different curvatures of the two halves of the TBP β-sheet determine directionality (34). Although Juo et al. (35) have suggested that proline 285 (human TBP numbering) determines the direction of TBP binding through its minor groove interaction with the first T-A base pair, this proposal is not supported by mutational analysis (31). Moreover, in contrast to the crystallographic studies, TBP in solution binds the TATA box in both orientations with only a slight preference for the consensus orientation (36), and there are no unfavorable interactions when TBP binds to the TATA box in the reverse orientation (37).

Notably, the deformation of the TATA element by TBP (32, 33, 38) creates a structurally invariant nucleoprotein complex that serves as the receptor for TFIIB and TFIID (39), which increase the preference for normal orientation of the TBP (40). Thus, although the polarity of transcription is independent of TATA box orientation (8), TBP probably binds in the normal orientation to the inverted TATA box in BSP because a TA-rich sequence with Ts predominant on the 5’ end appears to provide the necessary binding and orientation. When the flanking nucleotides are inverted a C is placed at position 7. Although crystallographic studies indicate that this substitution can be accommodated through the formation of a C:G Hoogsteen base pair (3), transcription is reduced (Fig. 4).

That the sequence of the TATA box can function well in either orientation or with various T → A mutations is consistent with the ability of TBP to recognize variants of nucleotides that form the minor groove of the TATA element, which is believed to be important in preventing potentially lethal mutations without perturbing subsequent steps in PIC assembly (35). However, this conflicts with the frequent conservation of gene specific TATA box sequences, such as the inverted sequence in the ICE/ICE-RICE-f5 elements relative to the PIC as deletions are made is shown in cross-section (C). The effect of reducing the distance between the ICE and TATA on transcription is depicted with bending of the DNA allowing interaction between the transcription factors associated with the two elements (A). As shown in B, for the f5-RICE-f5 to interact it must be flipped as well as rotated as deletions are made. As a result, in contrast to the ICE, transcriptional activity is not recovered after the deletion of 10 nucleotides (approximately one turn of the helix). The change in the orientation of the ICE/ICE-RICE-f5 elements relative to the PIC as deletions are made is shown in cross-section (C).

That immediate upstream activators may be involved in the directionality of TBP binding and transcription is indicated from the topographical analysis of a synthetic promoter in which an ATTTGCACT octamer was found to direct transcriptional polarity through a downstream TATA complex (8). More recently, the Gal4-VP16 and Gal4AH activators were shown to enhance the orientation and axial specificity of the TBP/TATA complex (40). Because the CCAAT-binding protein NF-Y has been shown to be essential for the recruitment of general transcription factors and RNAPII to the core promoter of the γ-globulin (42), OPN and ODF/RANKL genes, as well as to the TATA-less E2F1 gene promoter (43), it is conceivable that directionality of transcription could be determined by the proximal CCAAT element that is present in ~30% of eukaryotic genes (12). Accordingly, the orientation of the TATA sequence would be irrelevant for basal transcription.
The optimal transcription observed in the BSP promoter after 10 nucleotides were deleted suggests that, given the limited space remaining for NF-Y and TBP binding, NF-Y could be interacting directly with TBP and or PIC components. That NF-Y-mediated transcription is retained following the nucleotide deletions was confirmed by the positive and negative effects of NF-YA and NF-YA-DN expression vectors, respectively, in transcription assays and also by ChIP assays (Figs. 5 and 6). Thus, NF-YC/NF-YB has been shown to interact in vitro with TBP but not with a preformed TBP-TATA element (13, 44), which would be in agreement with a role in recruitment of the TBP (16). The NF-YB and NF-YC subunits include histone fold motifs required for dimerization to form a structure that is closely related to H2A/H2B and particularly to the NC2a/NC2b dimers (9). Thus, NF-Y subunits could activate transcription by directly interacting with TBP or by blocking the binding of the repressor NC2 to TBP (45). Alternatively, the NF-Y subunits may promote transactivation by interacting with complementary histone fold structures of TAFs, which are present in the TFIIID complex.

In contrast to the TATA box, reversing the ICE box resulted in the abrogation of transcriptional activity, whereas inclusion of flanking nucleotides recovered both NF-Y binding and transcription. However, with the inclusion of more than 8 flanking nucleotides, transcription was reduced despite strong NF-Y binding. Previous studies have demonstrated the importance of the CCAAT and TATA elements in the proximal promoters of the BSP gene that are relevant to a broad range of genes in which these elements are also involved. Previous studies have reported that transcriptional efficiency is unchanged when the distance between the upstream activator and TATA box was reduced from 24 to 5 nucleotides (8). In these studies the spacing between the upstream activator and TATA box was shown to be less constrained than between the TATA box and the transcription initiation (Inr) site, which would argue against a possible downstream repositioning of the PIC complex as the distance between the ICE and TATA boxes are reduced. However, although the transcriptional activity reported by Xu et al. (8) was unchanged until 3 nts separated the two elements, the cyclical changes observed in transcriptional activity in relation to the positioning of the ICE and TATA boxes are consistent with maximal expression obtained when two PEA3/EBS, Ets-binding sites, driving expression of a reporter gene functioning as a minimal transcriptional initiator element, were located on the same face of the DNA helix (49). The two sites could be separated by <2–3 helical turns but, as we found, transcription was markedly reduced when the elements were displaced by half a helical turn. The results of these studies are also consistent with the increased energy required to change the DNA structure to allow interactions between proteins in the CCAAT and TATA complexes as they are progressively displaced on the DNA helix (Fig. 8). A deletion strategy similar to the one that we have used has shown that activation by CRP (cAMP receptor protein) of Escherichia coli is strictly dependent upon the helical phase between the CRP and RNA polymerase-binding sites (50), whereas the importance of proper helical phasing that we have observed with the ICE and TATA boxes has been shown for two operators involved in the repression of the gal promoters by GalR and HU (51).

That changes in DNA structure can be altered to accommodate transcription factor interactions with the PIC complex has been demonstrated by Dion and Coulombe (52), who used site-specific protein-DNA photo-cross-linking to show that the transcriptional activator GAL4-VP16 interacts with a TFIIA-TFIIB complex assembled on the TATA element. Although the TBP-TFIIA-TFIIB promoter topology was not altered significantly by the interaction with DNA, similar to the requirements of the CCAAT and TATA box interactions in our study, looping and bending was required, and the efficiency of transcription was dependent upon the location and orientation of the GAL4-binding site relative to the TATA box. Notably, in our studies the recovery of transcriptional activity after 5 deletions was slower than the initial loss of transcriptional activity, which may be due to the additional clockwise twisting required for realigning the CCAAT and TATA elements or the greater energy required to twist the DNA in a counterclockwise direction.

The ability of the DNA to undergo marked structural changes to accommodate interactions between transcription factors binding to the ICE and TATA boxes is conceivable given the effects of transcription factor binding that has been shown from the crystal structure determination of TBP-DNA (32, 33) and from circular permutation analysis of NF-Y-DNA complexes (53, 54). On binding TBP a sharp 80° bend is induced on the 5′-side of the TATA box, and the minor groove is dramatically widened (32, 33). Similarly, the trimeric NF-Y complex binds to both the major and minor grooves with high specificity and avidity (47) causing the DNA to bend by 60–80°. In both cases the structural changes imposed on the DNA are thought to be a prerequisite for transcriptional activation. However, it is apparent from our transcriptional analyses that the limits of DNA bending are exceeded as the distance between the f5-RICE-f5 and TATA box is reduced (Fig. 7), because this also requires the reversed ICE box to flip over into the orientation of the wild type CCAAT element. Collectively, we believe these studies provide important insights into the stereochemical requirements of CCAAT and TATA elements in the proximal promoters of the BSP gene that are relevant to a broad range of genes in which these elements cooperate in the regulation of basal transcription.

Acknowledgments—Assistance with plasmid preparations and transcription assays was provided by Kam-Ling Yao. We also thank Dr. Roberto Mantovani (Università di Milano, Milano, Italy) for generously providing recombinant pET derivatives of the NF-YA and NF-YDN expression vectors for these studies, and Drs. Linda Penn and Harry Elsholtz (University of Toronto) for advice with the ChIP assays.

REFERENCES
1. Pugh, B. F. (1996) Curr. Opin. Cell Biol. 8, 303–311
2. Roeder, R. G. (2005) FEBS Lett. 579, 909–915
3. Patikoglou, G. A., Kim, J. L., Sun, L., Yang, S.-H., Kodadek, T., and Burley, S. K. (1999) Genes Dev. 13, 3217–3230
4. Nikolov, D. B., and Burley, S. K. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 15–22
Stereochemical Relationship between CCAAT and TATA Elements

5. Burley, S. K., and Roeder, R. G. (1996) *Ann. Rev. Biochem.* **65**, 769–799
6. Tan, S., and Richmond, T. J. (1998) *Curr. Opin. Struct. Biol.* **8**, 41–48
7. Burley, S. K., and Kamada, K. (2002) *Curr. Opin. Struct. Biol.* **12**, 225–230
8. Xu, L. C., Thali, M., and Schaffner, W. (1991) *Nucleic Acids Res.* **19**, 6699 – 6704
9. Romier, C., Cocchiarella, F., Mantovani, R., and Moras, D. (2003) *J. Biol. Chem.* **278**, 1336 – 1345
10. Mantovani, R. (1999) *Gene (Amst)* **239**, 15–27
11. Maity, S. N., and Crombrugghe, B. D. (1998) *Trends Biochem. Sci.* **23**, 174–178
12. Mantovani, R. (1998) *Nucleic Acids Res.* **26**, 1135–1143
13. Bellorini, M., Lee, D. K., Dantonel, J. C., Zemzoumi, K., Roeder, R. G., Tora, L., and Mantovani, R. (1997) *Nucleic Acids Res.* **25**, 2174–2181
14. Motta, M. C., Caretti, G., Badaracco, G. F., and Mantovani, R. (1999) *J. Biol. Chem.* **274**, 1326–1333
15. Silvio, A. D., Imbriano, C., and Mantovani, R. (1999) *Nucleic Acids Res.* **27**, 2578–2584
16. Frontini, M., Imbriano, C., diSilvio, A., Bell, B., Bogni, A., Romier, C., Moras, D., Tora, L., Davidson, I., and Mantovani, R. (2002) *J. Biol. Chem.* **277**, 5841–5848
17. Fisher, L. W., and Fedarko, N. S. (2003) *Connect Tissue Res.* **44**, (Suppl. 1) 33–40
18. Ganss, B., Kim, R. H., and Sodek, J. (1999) *Crit. Rev. Oral Biol. Med.* **10**, 79–98
19. Kim, R., Shapiro, H., Li, J., Wrana, J., and Sodek, J. (1994) *J. Biol. Chem.* **269**, 31–40
20. Li, J. J., Kim, R. H., and Sodek, J. (1995) *Biochem. J.* **309**, 33–40
21. Li, J. J., and Sodek, J. (1993) *Biochem. J.* **289**, 625–629
22. Sodek, J., Li, J. J., Kim, R. H., Ogata, Y., and Yamauchi, M. (1996) *Connect Tissue Res.* **35**, 23–31
23. Mantovani, R., Pessara, U., Tronche, F., Li, X. Y., Knapp, A. M., Pasquali, J. L., Benoist, C., and Mathis, D. (1992) *EMBO J.* **11**, 3315–3322
24. Kim, R. H., and Sodek, J. (1999) *Cancer Res.* **59**, 565–571
25. Man, D. Y., Baryte-Lovejoy, D., Ho, C. S., Watson, J. D., Stoianova, A., and Penn, L. Z. (2004) *Nucleic Acids Res.* **32**, 3462–3468
26. Su, M., Bansal, A. K., Mantovani, R., and Sodek, J. (2005) *J. Biol. Chem.* **280**, 38365–38375
27. Carcamo, J., Maldonado, E., Cortes, P., Ahn, M. H., Ha, I., Kasai, Y., Flint, J., and Reinberg, D. (1990) *Genes Dev.* **4**, 1611–1622
28. Kim, R., Li, J., Ogata, Y., Yamauchi, M., Freedman, L., and Sodek, J. (1996) *Biochem. J.* **318**, 219–226
29. O’Shea-Greenfield, A., and Smale, S. T. (1992) *J. Biol. Chem.* **267**, 6450
30. O’Shea-Greenfield, A., and Smale, S. T. (1992) *J. Biol. Chem.* **267**, 1391–1402
31. Spencer, J. V., and Arnold, K. M. (2002) *Mol. Cell. Biol.* **22**, 8744–8755
32. Kim, J., Nikolov, D. B., and Burley, S. K. (1993) *Nature* **365**, 520–527
33. Kim, Y., Geiger, J. H., Hahn, S., and Sigler, P. B. (1993) *Nature* **365**, 512–520
34. Suzuki, M., Allen, M. D., Yagi, N., and Finch, J. T. (1996) *Nucleic Acids Res.* **24**, 2767–2773
35. Jun, Z. S., Chiu, T. K., Leiberman, P. M., Baikalov, I., Berk, A. J., and Dickerson, R. E. (1996) *J. Mol. Biol.* **261**, 239–254
36. Cox, J. M., Hayward, M. M., Sanchez, J. F., Geggas, L. D., van der Zee, S., Dennis, J. H., Sigler, P. B., and Schejput, A. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 13475–13480
37. Malsiewicz, K., and Orntstein, R. L. (1996) *J. Biomol. Struct. Dyn.* **13**, 593–600
38. Kim, J. L., and Burley, S. K. (1994) *Nat. Struct. Biol.* **1**, 638–653
39. Geiger, J. H., Hahn, S., Lee, S., and Sigler, P. B. (1996) *Science* **272**, 830 – 836
40. Kays, A. R., and Schejput, A. (2000) *Chem. Biol.* **7**, 601–610
41. Tamura, T., Sumita, K., Fujino, I., Aoyama, A., Horikoshi, M., Hoffmann, A., Roeder, R. G., Muramatsu, M., and Mikoshiba, K. (1991) *Nucleic Acids Res.* **19**, 3861–3865
42. Duan, Z., Stamatoyannopoulos, G., and Li, Q. (2001) *Mol. Cell. Biol.* **21**, 3083–3095
43. Kabre, Y., Yamada, I., Uga, H., Yamaguchi, Y., Wada, T., and Handa, H. (2005) *Mol. Cell. Biol.* **25**, 512–522
44. Bellorini, M., Zemzoumi, K., Farina, A., Berthelsen, J., Piaggio, G., and Mantovani, R. (1997) *Gene (Amst)* **193**, 119–125
45. Kamada, K., Shu, F., Chen, H., Malik, S., Stelzer, G., Roeder, R. G., Meisterernst, M., and Burley, S. K. (2001) *Cell* **106**, 71–81
46. Bucher, P. (1990) *J. Mol. Biol.* **212**, 563–578
47. Bi, W., Wu, L., Crouzry, F., Crombrugghe, B. d., and Maity, S. N. (1997) *J. Biol. Chem.* **272**, 26562–26572
48. Liang, S. G., and Maity, S. N. (1998) *J. Biol. Chem.* **273**, 31590–31598
49. Yu, M., Yang, C. J., Schmidt, T., Chinenov, Y., Wang, R., and Martin, M. E. (1997) *J. Biol. Chem.* **272**, 29060–29067
50. Ushida, C., and Aiba, H. (1990) *Nucleic Acids Res.* **18**, 6325–6330
51. Lewis, D. E., and Ahl, S. (2000) *J. Biol. Chem.* **275**, 2497–2504
52. Dion, V., and Coulombre, B. (2003) *J. Biol. Chem.* **278**, 11495–11501
53. Liberati, C., Sgarra, R., Manfioletti, G., and Mantovani, R. (1998) *FEBS Lett.* **433**, 174–178
54. Ronchi, A., Bellorini, M., Mongelli, N., and Mantovani, R. (1995) *Nucleic Acids Res.* **23**, 4565–4572