Purification and Characterization of a Protein Binding to the SP6 k Promoter

A POTENTIAL ROLE FOR CArG-BOX BINDING FACTOR-A IN k TRANSSCRIPTION*

(Received for publication, May 1, 1998)

Mats Bemark‡, Henric Olsson§, Dick Heinégård§, and Tomas Leandersson†¶

From the ‡Immunology and §Connective Tissue Biology Groups, Department of Cellular and Molecular Biology, Lund University, S-220 07 Lund, Sweden

A protein interacting with an A-T-rich region that is a positive control element within the SP6 k promoter was purified and identified as CArG-box binding factor-A. The purified protein was shown to interact specifically with the coding strand of single-stranded DNA and, with lower affinity, with double-stranded DNA. A mutation that inhibited binding of the protein to the A-T-rich region also aborted the transcriptional stimulatory effect of the region. Two Ets proteins, PU.1 and elf-1, that have previously been shown to bind to an adjacent DNA element were shown to physically interact with CArG-box binding factor-A. An antiserum raised against the protein recognized two different forms indicating either that different splice-forms of CArG-box binding factor-A are expressed, or that the protein is subject to post-translational modification.

The control of transcriptional initiation in eucaryotes is a multilayered process. At one end of the spectrum are the very specific interactions between distinct sequence motifs and defined transcription factors that can interact with a specific sequence only. Secondary protein-protein interactions between such DNA-binding factors and so-called transcriptional adapter molecules also show molecular specificity (1–4) and can easily be envisioned to have a defined role in the control of expression of a given gene. The mechanism of action and the specificity of other control elements pertinent to gene expression are of more general nature. Here, one may distinguish between control mechanisms of gene expression from a given locus and those that are acting downstream of this decisive event to facilitate efficient transcriptional initiation and elongation induced by the sequence-specific transcription factors mentioned above. Among the former are locus control regions (5), matrix attachment regions (6), histone acetylation (7), and the action of the SWI/SNF complex (8). The latter include specific DNA bending proteins (9), topoisomerases (10), stress induction by the matrix (10) as well as specific single-strand binding proteins (11–13). Additional biochemical events that modulate the efficacy of gene expression are hairpin extrusion (14), DNA melting (15–17), Z-DNA formation (18), and supercoiling due to torsional stress induced by the transcription machinery (19, 20).

During recent years, several proteins have been described that interact with single-stranded promoter elements and that appear to influence transcriptional initiation (11, 12, 21–27). These proteins usually show sequence specificity with regard to binding, although they are probably not as restricted as double-stranded DNA binding transcription factors (28). Several of these single-stranded DNA-binding proteins are identical or related to proteins that are found within the heterogeneous nuclear ribonucleoprotein (hnRNP)1 complex (11, 12, 23–25) that is formed on heterogeneous nuclear RNA and is thought to be important for protection, splicing, and transport of RNA (29). Thus, it is feasible that these proteins have dual functions, as have been proposed for other DNA/RNA-binding proteins in prokaryotes and in polymerase III genes (30, 31).

The transcriptional regulation of the immunoglobulin (Ig) genes is complex. Although each rearranged gene has a distinct promoter, distal enhancers are found in introns and 3′ of the genes (32). Within these enhancers, a variety of DNA elements involved in the transcriptional control have been defined (32). Furthermore, close to the intron enhancers, matrix attachment regions are found, and recent evidence indicates that these function as locus control regions (33, 34).

Ig k promoters show sequence conservation within but not between V gene subgroups (35, 36). Detailed functional studies have shown that k promoters contain several DNA elements involved in transcriptional regulation (37–45). The octamer, which is a binding site for the ubiquitous OCT1 and the B cell-specific OCT2 transcription factors, is a key control element in k promoters and when mutated the promoter is inactivated (37). However, the promoter is dependent on other, octamer-dependent, transcriptional control elements for proper function (37). Such elements have been identified both 5′ and 3′ of the octamer and examples thereof are the κ-Y element (44), the pentadecamer (pd) element containing an E-box of the E2A type (-CAGNTG-) (38, 41), and the CCCT element (39, 42). All three elements interact with distinct proteins in electrophoretic mobility shift assay (EMSA) and interact functionally with the octamer (42, 43). So far, only the proteins that interact with the κ-Y element have been identified, PU.1 and elf-1, both of which belong to the Ets family of transcription factors (43, 45).

In this study, we purified a protein that interacts with a functionally important A-T-rich region within the SP6 k pro-

* This work was supported by grants from the Swedish Cancer Society, the Swedish Medical Research Council, Kocks Foundation, the Osterlund Foundation, the Crafoord Foundation, and the LEO Research Foundation. 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.

† To whom correspondence should be addressed: Immunology Group P.O. Box 7031, S-220 07 Lund, Sweden. Tel.: 46-46-2224160; Fax: 46-46-2224218; E-mail: Tomas.Leandersson@immuno.lu.se.

1 The abbreviations used are: hnRNP, heterogeneous nuclear ribonucleoprotein; PAGF, polyacrylamide gel electrophoresis; GST, glutathione S-transferase; CAT, chloramphenicol acetyltransferase; EMSA, electrophoretic mobility shift assay; pd, pentadecamer; pdLMW, pd element low molecular weight shift; pdMMW, pd element medium molecular weight shift.
mM phenylmethylsulfonyl fluoride, 20% glycerol, and 0.1% Nonidet (Amersham Pharmacia Biotech). The protein was cleaved by incubating with 0.5 units of thrombin (Sigma) in 200 mM digestion buffer (20 mM Tris (pH 8.0), 25 mM KCl, 5 mM MgCl₂, 1 mM dithiothreitol, and 0.1% Nonidet P-40) (48). The material that precipitated between 70 and 90% ammonium sulfate saturation was dialyzed against buffer Z and was loaded onto a native DNA cellulose column (Amersham Pharmacia Biotech). The protein was further purified on a MonoQ™ column (Amersham Pharmacia Biotech). The protein was purified using a QuickPrep™ micro mRNA purification kit (Amersham Pharmacia Biotech) according to the manufacturer’s instructions (Amersham Pharmacia Biotech), whereafter it was further purified on a MonoQ™ column (Amersham Pharmacia Biotech) column equilibrated in buffer Z. After washing, bound proteins were eluted stepwise in 100 mM steps up to 1 M KCl in buffer Z. Fractions containing the pdLMW binding activity (41) (500 mM and 600 mM KCl) were pooled and dialyzed against buffer A (20 mM Tris (pH 8.0), 25 mM KCl, 5 mM MgCl₂, 1 mM dithiothreitol, and 0.1 mM phenylmethylsulfonyl fluoride) and loaded onto a MonoQ™ column (Amersham Pharmacia Biotech) column equilibrated in buffer A. After washing with buffer A, bound proteins were eluted with a continuous gradient of buffer B (as A but 500 mM KCl).

SDS-PAGE, Western Blots, Trypsin Digestion, Amino Acid Sequencing, and Northern Blots—SDS-PAGE and electroblotting were performed according to standard procedures (49), and blotted proteins were immunodetected by chemiluminescence (50). For blocking of the antiserum, 3 μl of protein G purified sera was preincubated with 50 μg of recombinant protein for 30 min on ice before it was added to the membranes. For amino acid sequencing, the protein band was identified in the gel after Coomassie Blue staining, excised, and sent to SLU (Bo Ek, Uppsala, Sweden) for trypsin digestion, peptide purification by reversed phase high performance liquid chromatography, and amino acid sequencing. Northern blots were performed according to standard procedures using 0.5 μg of poly(A) RNA per lane (49). The RNA was purified using a QuickPrep™ micro mRNA purification kit (Amersham Pharmacia Biotech).

**Cloning of CArG-box Binding Factor-A, Expression in Bacteria, and Production of Polyclonal Serum—CArG-box binding factor-A was cloned from a HybriZAP cDNA library made from murine B cells that had been stimulated with lipopolysaccharide for 72 h, according to the manufacturer’s instructions (Stratagene), using an end-labeled oligonucleotide (5’-ATCAAGGTTGCCCAGCCCAAAGAG-3’) according to standard procedures (49). A clone that contained the complete 5’ sequence but lacking a part of the 3’ untranslated region was cloned (23).

The cDNA was cut with NcoI, blunt-ended, and then cut with SacI and inserted between the Smal and the SacI site in the pGEM3Z plasmid (Promega). This plasmid was completely digested with EcoRI and then partially digested with BamHI, and the CArG-box binding factor-A fragment derived in this manner was inserted in frame with the glutathione S-transferase gene between the BamHI and the EcoRI sites in the pGEX-2 vector (Amersham Pharmacia Biotech). The protein was expressed and purified on glutathione-Sepharose according to the manufacturer’s instructions (Amersham Pharmacia Biotech), whereas it was further purified on a MonoQ™ column (Amersham Pharmacia Biotech). The protein was cleaved by incubating with 0.5 units of thrombin (Sigma) in 200 μl of digestion buffer (20 mM Tris (pH 8.0), 150 mM NaCl, 2.5 mM CaCl₂, and 10% glycerol) at room temperature for 2–3 h.

The polyclonal antiserum was raised in rabbits according to standard procedures using the GST-CArG-box binding factor-A fusion protein. The IgG fraction was purified on a Protein G HiTrap™ column (Amersham Pharmacia Biotech) according to the manufacturer’s instructions and dialyzed against phosphate-buffered saline.

**The Deletion Mutants—**Variants 3, 7, 8, 9, and 10 were generated by PCR (15 cycles) using primers 1 and 5, 6, 7, or 8, respectively, and the cloned CBF-A cDNA as a template. Variants 4, 5, and 6 were generated by primary PCRs using primer 5 and 6, 7, or 5, respectively, or primer 1 and 4 for 15 cycles to generate fragment 1–4. The fragments were gel-purified, pre-PCR in the absence of primers was performed with fragment 4 and 1, 2, or 3 for 20 cycles, new Taq enzyme and primers 1 and 5 were added to the reactions, and 15 cycles more were performed. After digestion with BglII and EcoRI, the fragments were then cloned between the BamHI and EcoRI sites in the GST vector pGEX-2 (Amersham Pharmacia Biotech). All constructs were

![Figure 1](image-url)
verified by sequencing of the first 250 base pairs using the pGEX 5′ sequencing primer (Amersham Pharmacia Biotech). The sequence of the primers used were: 1) 5′-GGAGAGATCTCCGCCGAGGGAAACCAGA-3′; 2) 5′-GGAGAGATCTCCGCCGAGGGAAACCAGA-3′; 3) 5′-GGAGAGATCTCCGCCGAGGGAAACCAGA-3′; 4) 5′-GGAGAGATCTCCGCCGAGGGAAACCAGA-3′; 5) 5′-GGAGAGATCTCCGCCGAGGGAAACCAGA-3′; 6) 5′-GGAGAGATCTCCGCCGAGGGAAACCAGA-3′; 7) 5′-GGAGAGATCTCCGCCGAGGGAAACCAGA-3′; 8) 5′-GGAGAGATCTCCGCCGAGGGAAACCAGA-3′.

The deletion mutants were expressed in bacteria and purified on glutathione beads as described by the manufacturer (Amersham Pharmacia Biotech), and subsequently dialyzed against phosphate-buffered saline for 8 h with one buffer exchange. The amounts of the different variants were estimated by A280 measurement.

EMSA, Transfections, and CAT Assays—EMSA, transfections, and CAT assays were performed as described (43), but for the EMSAs single-stranded probe was used unless stated different in the figure legends. Preimmune or immune serum was added before the addition of the probe. The reactions were then preincubated for 30 min at room temperature. The sequences of the probes can be found in the figures.

For the EMSA in Fig. 5B, the promoter region of the SP6 κ gene was cloned between the PstI and the HindIII sites of the pGEM3Z vector (Promega). An end-labeled promoter probe was generated by Klenow fill-in labeling after XhoI digestion followed by digestion with HindIII. Native PAGE purified probe in bind buffer was denatured at 95°C for 5 min and then rapidly cooled on ice. 6000 cpm of native or denatured probe was added to 25 μl of footprint binding buffer with different amounts of protein as indicated. The reactions were incubated for 5 min at 37°C, whereafter an aliquot of 10 μl was separated on a 5% native PAGE gel in 1× TBE. Stop buffer (0.67% SDS, 30% glycerol, 0.3 mg/ml tRNA, 0.3 mg/ml proteinase K, and 0.3% bromphenol blue) (51) was then added, and the reaction was continued for 5 min. The reactions were then separated by native PAGE as above.

DNase Footprinting Assays and Methylation Interference—The footprints were performed according to standard procedures, and single-stranded or double-stranded probes (20 000 cpm) were used as indicated. After preincubation for 30 min in 50 μl of bind buffer (10 mM HEPES (pH 7.9), 50 mM KCl, 5 mM MgCl2, 0.5 mM EDTA, 0.5 mM spermidine, 0.5 mM dithiothreitol, and 10% glycerol) on ice, 50 μl of 5 mM CaCl2 and 10 mM MgCl2 was added and finally 0.25 μg of DNase was added. After incubation for 1 min the reaction was stopped by the addition of 100 μl of DNase stop buffer (200 mM NaCl, 20 mM EDTA, 1% SDS, 250 μg/ml yeast tRNA), whereafter the reaction was phenol-extracted and ethanol-precipitated. The reactions were separated on 20% denaturing PAGE gels. The ladders were made according to the rapid protocol for G/A Maxam-Gilbert sequencing (49). The methylation interference assay was performed essentially as described in Ref. 41, although a single-stranded probe was used and that the probe was gel-purified after methylation. Recombinant thrombin-cut CBF was used in EMSA with methylated probe (100,000 cpm), and about 50% of the probe was shifted. The bound and free probe was identified after autoradiography, excised and eluted over night. The reaction was separated on a 15% denaturing PAGE gel.

CAR G-box Binding Factor A Pull-down Assay—Recombinant GST or GST/CAR G-box binding factor fusion protein was bound to glutathione-Sepharose (Amersham Pharmacia Biotech) in NETN buffer (20 mM Tris,

![Graph](image_url)
RESULTS

**Functional Activity and Protein Interactions with an A-T-rich Region in the SP6 κ Promoter pd Element**—The region 5′ of the octamer in the murine SP6 κ promoter has two sequence elements, the pd and κ-Y elements, that stimulate transcription by synergistic interactions with the octamer (41, 43–45). The pd element can be further divided into two sites, a 5′ E-box of the E2A type and a 3′ A-T-rich region (41). The individual function of these three sites was tested by transient transfections of CAT reporter constructs into lipopolysaccharide-stimulated B cells (41). A promoter fragment containing the pd, the κ-Y, and the octamer element stimulated transcriptional initiation when placed upstream from a TATA-box, as compared with the TATA-box alone (Fig. 1A; compare the first two lanes). When the octamer was mutated only limited transcriptional stimulation was observed (third lane); this confirmed earlier studies showing that the octamer is obligate for promoter function (37). The individual function of the three sites 5′ of the octamer was tested by introducing mutations in the reporter construct. Deletion of the 5′ E-box of the pd element, or mutations in either the A-T-rich region or the κ-Y element all diminished the costimulatory function (fourth through sixth lanes). Thus, the interdependence of the three 5′ elements for octamer costimulation was confirmed (43).

Next, the pd element was used as probe in an EMSA with nuclear extract from the plasmacytoma cell line J558L as protein source (Fig. 1B). As described previously, two complexes were formed (pdMMW and pdLMW; Ref. 41). When the A-T-rich region was mutated, no pdLMW complex was formed while the pdMMW complex was unaffected. A mutation of the E-box did not affect the pdLMW complex but the pdMMW complex was reduced in intensity. The binding of the pdMMW complex merits some comment, as it may seem to be in contrast to an earlier report where mutations of the E-box disrupted binding (41). However, different methods were used to generate the nuclear extracts; the method of Schreiber et al. (52) was used in the earlier report, whereas a modification of the Dignam method (47) was used here. Thus, the pdMMW shift here appears to be composed of two overlapping complexes: a specific E-box binding complex that does not bind to the mutated E-box and a nonspecific binding complex that also interacts with single-stranded DNA (data not shown).

**Purification and Amino Acid Sequencing of the pd A-T Region Interacting Protein**—To identify the A-T region interacting protein, we developed a three-step method that was used for purification (Fig. 2A). Nuclear extract was made from the J558L plasmacytoma cell line according to a modification of the original Dignam method (46, 47), and the extract was sequentially precipitated with ammonium sulfate in 10% steps. The pdLMW activity precipitated mainly in the 70–80% and 80–90% ammonium sulfate saturation fractions, while the pdMMW activity (and most other transcription factors) precipitated at concentrations below 40% saturation (data not shown). The 70–80% and 80–90% fractions were pooled and loaded onto a native DNA cellulose column, and while most contaminating proteins passed through the column at 100 mM salt concentration, the pdLMW activity was eluted in the 500 and 600 mM fractions (data not shown). These two fractions were pooled, loaded onto a MonoQ™ ion exchange column, and eluted with a continuous gradient of KCl. Every second fraction was tested for protein content by SDS-PAGE and for pdLMW binding activity in EMSA (Fig. 2B, left panels). The activity eluted in one peak in fractions 24–26, as did an approximately 35-kDa protein, a molecular mass in agreement with a preliminary characterization of the pdLMW activity (41). A contaminating protein of higher molecular mass eluted earlier than and overlapping with the 35-kDa protein. The purification steps were monitored by SDS-PAGE (Fig. 2B, right panels). Although some contaminating proteins were evident after the DNA-cellulose step, only the higher molecular mass protein contaminated the purified protein after the MonoQ™ step. Hence, we conclude that a 35-kDa protein co-purified with the pdLMW forming activity during the three purification steps.

To identify the protein, fraction 25 was acetone-precipitated and separated on SDS-PAGE. After staining with Coomassie Blue to identify the position of the protein, the band was excised from the gel, trypsin digested, and the resulting peptides were separated using reverse phase high performance liquid chromatography. Two peptides were amino acid sequenced, and the sequences were used to search the EMBL database using the FASTA algorithm. CArG-box binding factor-A (CBF-A; Ref. 23) was the highest scoring candidate for both peptides (Table I; only one mismatch was observed at an
ambiguous amino acid position), and this identification was further strengthened by the fact that CBF-A has potential trypsin digestion sites just upstream from either of the sequenced peptides.

CBF-A Is Part of the pdLMW Complex—The cDNA for CBF-A was cloned from a library derived from murine B cells that had been stimulated with lipopolysaccharide for 72 h, using an oligonucleotide probe corresponding to nucleotides 883–907 in the published sequence (23). A GST fusion protein was expressed in *Escherichia coli* by inserting the coding region from amino acid 10 to the stop codon in the pGEX-2T expression vector and purified as described under “Experimental Procedures.” Subsequently, the purified protein was tested in EMSA for interactions with the pd probe. Although purified GST did not form a complex with the probe in EMSA, purified GST-CBF fusion formed two complexes, a stronger with lower mobility and a weaker with somewhat higher mobility (Fig. 3A, left panel, lanes 1 and 2). The same complexes were formed even after MonoQ™ purification (lane 3), and the complex with higher mobility appears to be a breakdown product of the GST-CBF fusion protein, as a minor protein band is evident in SDS-PAGE with a slightly lower molecular weight than the major band (data not shown). When the fusion protein was cleaved by thrombin the formation of both complexes were diminished and a novel strong complex was formed with approximately the same mobility as the purified protein (Fig. 3A, right panel). As the same amount of protein was loaded in lanes 1 and 2, it appeared that the GST part of the fusion protein

CBF-A and κ Transcription

FIG. 4. A, Western lots were performed with preimmune or immune sera using uninduced bacterial extract (left lanes), expressed GST (middle lanes), or GST-CBF (right lanes). B, Western lots were performed as in A but the immune serum was blocked with GST (middle panel) or GST-CBF (right panel) before it was incubated with the blot. C, Western lots were performed with J558 extract, S194 extract or purified CBF using either preimmune serum (left panel) or immune serum (middle panels). To the right the preimmune serum was blocked as in B before it was used. D, Northern blots are shown where poly(A)+ mRNA from the indicated B cell lines were probed with a CBF-A probe and a parallel blot where the same mRNA samples were probed with a GADPH probe.

FIG. 5. A, left panel shows EMSAs with purified protein and different pd element probes as indicated. The probes were either from the coding strand (c) or non-coding (nc) of the promoter and were single-stranded (ss) or double-stranded (ds) as indicated. The binding to each to the probes was tested in EMSA as indicated with the same amount of purified protein in each lane (fraction 24) as source of protein. To the right, the binding of the purified protein was competed with either of the strands, double-stranded competitor or a single-stranded, unrelated competitor. B, an EMSA with double-stranded or heat-denatured full-length SP6 κ promoter as probe and purified CBF-A as source of protein is shown at the top, and an assay to determine of the probe is double-stranded or single-stranded is shown at the bottom. The experimental protocol is shown to the right.
disturbed binding. We conclude that recombinant CBF-A interacts with the pd probe in EMSA and has a mobility similar to that for the purified pdLMW protein.

The GST-CBF fusion protein was used to generate a polyclonal antiserum. The specificity was confirmed by Western blotting using recombinant protein and either preimmune or immune serum and blocking with the recombinant protein (Fig. 4; see below). Neither the preimmune nor the immune serum affected the pd probe alone in EMSA (Fig. 3 B, left panel, lanes 1–5), and while the preimmune serum did not affect the binding of the purified protein in EMSA, the immune serum did (Fig. 3 B, middle panel). The same result was obtained using the recombinant protein in EMSA (data not shown). Thus, the purified protein is immunologically cross-reactive with CBF-A. Next, we tested if the immune serum did interfere with the formation of the pdLMW complex in EMSA using nuclear extract as a source of protein. This was the case; although no effect on the pdLMW complex formation was evident when the nuclear extract was preincubated with the preimmune serum, it was diminished when immune serum was used (Fig. 3 B, right panel). However, some residual complex formation was observed that migrated at the slower front of the pdLMW complex. Note that the nuclear extract used in this experiment was made according to Schreiber et al. (52), since the pdLMW complex was first described using this extract (41), and hence the pdMMW like complex that was observed using the modified Dignam nuclear extract did not form as discussed above. However, the residual pd complex in the presence of the immune serum was also evident when a nuclear extract according to Dignam (47) was used as a source of protein (data not shown). Thus, although CARG-box binding factor-A is the main protein involved in the formation of the pdLMW complex, another minor protein complex with similar mobility can interact with the pd element.

Two Forms of CBF-A Can Be Detected in Nuclear Extracts—
The specificity of the antiserum was tested in Western blots (Fig. 4 A). The antiserum, but not the preimmune serum, did recognize either recombinant GST-CBF fusion protein or GST alone. No other bacterial proteins where detected in the assay, and while the interaction to GST could be blocked by either GST or GST-CBF, the interaction to GST-CBF could only be blocked by recombinant GST-CBF (Fig. 4 B).
Next, a Western blot with nuclear extract from J558 cells was probed with antiserum or preimmune serum (Fig. 4C). Although no interactions were evident using the preimmune serum, two protein bands were detected using the CBF antiserum. The interaction with the two protein bands could not be blocked by recombinant GST, whereas it was by the GST-CBF fusion protein (Fig. 4C, right panels). Surprisingly, two novel bands reproducibly appeared in the assay after blocking with recombinant GST-CBF. Our interpretation of this finding is that the CBF used as competitor interacted directly with other proteins in the nuclear extract, and that these interactions were subsequently detected by the anti-CBF serum. The mobilities of the two bands detected in nuclear extracts were subsequently compared with the mobility of purified CBF-A (Fig. 4C, left panel). The purified protein was detected as a single band that corresponded in mobility to the faster moving of the two bands in the nuclear extract. To test if the slower mobility protein was specific to the J558 cell line, the S194 myeloma cell line was also tested. As shown in Fig. 4C (middle panel), both bands were present also in this cell line. The same was also true for a panel of different murine B cell lines (data not shown).

To investigate whether CBF-A was encoded by multiple mRNA species we performed a Northern blot using poly(A)+ RNA from three mouse B cell lines (Fig. 4D). All expressed CBF-A RNA, but a weakly hybridizing RNA with lower mobility could also be detected. The relative amount of this larger transcript, as compared with the major hybridizing band, did not correspond to the relative amounts of the two different proteins that were observed in Western blots. Rather, the observed pattern is consistent with the higher molecular weight RNA being a splicing intermediate. Hence, we prefer to interpret our data as to indicate that CBF-A undergoes post-translational modifications. However, the resolution of our Northern blot makes it impossible to exclude that more than one splice variant of CBF-A were expressed if these were of a similar size (60).

**CBF-A Interacts with Either Single- or Double-stranded DNA**—During the purification of the A-T-rich region interacting protein, it was noted that the pdLMW activity copurified with a single-strand binding activity. Therefore, the binding of the protein was further characterized in EMSA with the pd element labeled on either strand, in single- or double-stranded form, as probe. The protein interacted strongly with single-stranded probe labeled on the coding strand, while less binding was evident using double-stranded probe (Fig. 5A, left panel, lanes 1 and 2). Even though the protein interacted with the noncoding strand forming a complex of similar mobility (lanes 3 and 4), this binding was considerably weaker than the binding to the coding strand. However, both double-stranded DNA or either of the strands competed for CBF-A binding to a coding strand probe, while an unrelated single-stranded oligonucleotide did not (Fig. 5A, right panel).

It could not be excluded that a significant fraction of the double-stranded probes and competitors used were in fact single-stranded, or that the two strands were separated before binding. We therefore developed a method that directly addressed if the protein could bind double-stranded DNA, as the issue whether or not CBF-A could also interact with double-stranded DNA was important for our understanding of its biological activity. The SP6 k promoter was inserted into the pGEM3Z plasmid and end-labeled as described under “Experimental Procedures.” Then either native double-stranded probe or heat-denatured probe was incubated with increasing amounts of purified protein. After a 5-min incubation, half of each reaction was loaded onto a non-denaturing PAGE, while SDS-containing stop buffer was added to the remainder of the reactions and they were then incubated for another 5 min before loading. A mobility shift representing binding of protein to the probe was observed with either double- or single-stranded probe, although a smaller amount of protein was needed to shift the single-stranded probe (Fig. 5B, upper panel). The shift of double-stranded probe was not due to the probe being denatured, as judged by the difference in mobility that was evident between single-stranded and double-stranded probe after the addition of SDS stop buffer (Fig. 5B, lower panel). Rather, it appeared that the single-stranded probe was reannealed during the incubation with protein. Thus, CBF-A is able to interact with double-stranded DNA but with a lower efficiency than to a single-stranded template.

**Mapping of the Interaction between CBF-A and DNA**—We next determined which part of the pd element that was contacted by CBF-A. A single-stranded probe was used that corresponded to the coding strand of the promoter and included the pd, the \( \kappa-Y \) and the octamer elements. When increasing amounts of purified protein or recombinant CBF-A were added to the probe before the addition of DNase, a fully protected footprint that extended over the A-T-rich region of the pd element was observed (Fig. 6, A and B). Partial protections of the k-Y element proximal to the pd element and of the pd E-box were also evident, while the DNase reactivity of the rest of the k-Y element and the octamer was enhanced. A similar footprint over the A-T-rich region was observed when the pd element only was used as probe (data not shown), showing that the single-stranded footprint was specific for the A-T-rich region. The interaction between single-stranded DNA and CBF-A was further analyzed in a methylation interference assay using recombinant CBF-A (Fig. 6C). In this assay, two A residues were important for CBF interaction (marked with arrowheads), and these were identical to those shown to be critical for function in Fig. 1. A partial protection of the last G residue of the pd
element E-box was also detected (marked with an asterisk; the results of the footprints and the methylation interference assays are summarized in Fig. 7). These results were those observed previously in a methylation interference using nuclear extracts as a source of protein (41). We also attempted footprint and methylation assays of double-stranded probes. However, in these assays, we could only detect very weak protection at high protein concentrations (Fig. 6D and data not shown). Thus, whereas a specific interaction could be shown between single-stranded DNA and CBF-A, the interaction between the protein and double-stranded DNA was either unspecific or of low affinity to be visualized by these techniques. This conclusion is also supported by the data shown in Fig. 5.

The specificity between CBF-A and single-stranded DNA was further analyzed in EMSA using single-stranded probes with different mutations as compared with the wild type sequence and recombinant or purified CBF-A (Fig. 7A). The wild type sequence interacted with either of the proteins (lane 1, pd), as did probes with mutations in the E-box or in the core of the κ-Y element (lanes 3 and 6, mut1 and mut2). Probes with mutations of the 3' part of the E-box (lane 4, mut2) or the two functionally important A residues that were protected in methylation interference (lane 2, ATmut) did not interact with the proteins, while a weak binding was evident using a probe with three T residues mutated downstream of the A residues (lane 5, mut 3). These results were in good agreement with the results of the footprint and methylation interference analysis and are summarized in Fig. 7B.

We next expressed a set of deletion mutants of the CBF-A protein in bacteria and purified them on glutathione-Sepharose beads as described under “Experimental Procedures” (schematically shown in Fig. 8B). All the variants were expressed and yielded recombinant proteins of the expected sizes, and no extensive breakdown was evident in Western blots (Fig. 8A, right panel). The deletion mutants were then tested for DNA interactions in EMSA using the single-stranded pd element as a probe (Fig. 8A, left panel). All variants that had two RNP domains intact interacted strongly with DNA (lanes 2, 3, 7, and 9), and a weak interaction was also evident for two variants that contained one RNP domain (lanes 4 and 5). Neither the GST tag alone (lane 1) nor a mutant where both RNP domains had been deleted (lane 6) showed significant DNA binding. Two of the recombinant proteins contained one RNP domain but did not bind DNA (lanes 8 and 10). In these mutants, the RNP domain was directly fused to the GST tag, which might interfere with DNA binding. In conclusion, the RNP domains of CBF-A were involved in the interaction with DNA and both of them were needed for a high affinity interaction.

CBF-A Interacts Directly with the Ets Proteins PU.1 and elf-1—The close proximity between the CBF-A binding site and the ets site in the SP6 κ promoter made us assay for interactions between their ligands. As shown in Fig. 9A, recombinant GST/CBF-A fusion protein bound to glutathione beads was able to pull down about 3.5% of in vitro translated PU.1 or elf-1, whereas GST alone or empty beads did not. The specificity of the interaction was demonstrated by that neither in vitro translated OCT2 nor E47/E12 protein were pulled down by GST/CBF-A (data not shown and Fig. 9A). Thus, CBF-A can interact with the Ets proteins elf-1 and PU.1 in vitro.

To map the interaction domain of CBF-A with Ets proteins, the deletion mutants that were used in Fig. 8 were tested for interactions with PU.1 (schematically shown in Fig. 9D). Fig. 9B shows a typical experiment. All the deletion mutants interacted with PU.1 and pulled down about 50% of the amount that was pulled down by the full-length variant used in Fig. 9A, with the exception of variant 7 (lane 10) that only contained one RNP domain. The same result was obtained using elf-1 instead of PU.1 (data not shown). To verify that all the GST fusion proteins bound to the glutathione beads and were not broken down during the incubation with reticulocyte lysate, we made a Western blot of recombinant protein eluted from glutathione beads after incubation with reticulocyte lysate. All the proteins were detected in the assay, but variant 4 was partially broken down and lower amounts were bound of variants 3, 5, and 7. If the amount of PU.1 pulled down with these deletion mutants was corrected for the lower amount of CBF-A recovered, it may be concluded that all mutants except 10 were equally efficient in the assay. Thus, the interaction between PU.1 and CBF-A is complex and involves two separate interaction domains on CBF-A; the RNP domains and the carboxyl-terminal region of

![Fig. 8. A, EMSAs with different deletion mutants of CBF-A and the single-stranded wild type pd element as probe are shown to the left (the probe has been run out of the gel). To the right is shown a Western blot of the same amounts of the different deletion mutants as were used in the EMSA. B, a schematic figure of the different deletion mutants used in A and their DNA binding characteristics is shown.](image-url)
the molecule. Whether these interactions are mutually exclusive or act in consort is currently unclear.

DISCUSSION

The A-T-rich region of the pd element has been shown to be active in octamer-dependent stimulation of \( \kappa \) transcription and to interact specifically with protein (pdLMW) in EMSA (41). Here we identify the pdLMW protein as CArG-box binding factor-A (CBF-A; Ref. 23). This protein was originally cloned from a \( \lambda \) expression library due to its interaction with the intron CArG-box of the human smooth muscle \( \alpha \)-actin gene, and it has high homology to D-box binding factor, a chicken DNA interacting protein cloned using the same approach (25). Both of these proteins were shown to interact mainly with single-stranded DNA, but in neither of the reports it was possible to exclude binding to double-stranded DNA. In this study, it was shown that the protein can interact also with double-stranded DNA or possibly with single-stranded structures that are formed within the double-stranded DNA; the protein was purified on a native DNA column and interacted with a double-stranded probe in EMSA, although with a lower efficiency than with single-stranded DNA.

CBF-A and the above mentioned D-box binding factor have RNP domains in common with hnRNP A and C proteins (29), a domain that is thought to be involved in the binding to nucleic acid (53). hnRNP A and C are parts of the hnRNP complex, which is thought to have a role in mRNA transport and splicing (29). However, some of the proteins that are associated with mRNA in hnRNP complexes have also been implicated as transcription factors. These include the hnRNP K protein, which associates with a region that is found in single-stranded form in the \( c\text{-myc} \) promoter and has potential as a transcriptional activator both \( \text{in vitro} \) and \( \text{in vivo} \) (11–13), and a protein related to hnRNP C that stimulates transcription of a viral gene upon overexpression (24). Furthermore, the TLS/FUS oncoprotein has been found in association with the hnRNP complex (54), while recent data have indicated a role for it in transcription as it is found associated with TBP in a subpopulation of TFIID complexes (55). Thus, many of the hnRNP proteins may have dual functions in the cell, \( i.e. \) they are involved in both splicing and transporting of RNA and in transcriptional control.

In this context, it should also be noted that pdLMW is located in the cytoplasmic fraction of resting B lymphocytes while being translocated to the nucleus upon mitogenic activation (41). This could indicate dual functions, but also be a regulatory step where the need for the protein in the nucleus is enhanced in activated cells. The antiserum that was raised against the protein interacted with two distinct bands in a Western blot, while only one major RNA was evident in Northern blot. Although it cannot be formally excluded that the two forms in the Western blot were due to cross-reactivity, or that two overlapping RNA splice forms of CBF-A exist (60), it is to our mind more likely that the two protein species arise due to post-translational modifications. These modifications could regulate either the localization or the function of the protein. The protein has several potential phosphorylation sites, and only the lower mobility form was evident during the purification. Thus,
post-translational modifications may interfere with the binding to nucleic acids.

CBF-A does not appear to be a classical transcription factor containing an activation domain since a multimerized pd element is a poor activator of transcription (41). Hence, we speculate that the protein rather has a structural role in the promoter, distorting the DNA either by bending or inducing other structural alterations that may facilitate efficient transcriptional activation. Furthermore, if the interactions with the Ets proteins PU.1 and elf-1 are to take place in vivo, CBF-A might act by increasing the local concentration of these transcription factors. CBF-A was originally described as a repressor of transcriptional activation. Furthermore, if the interactions with the Ets factors CBF-A was originally described as a repressor of transcriptional activation. Furthermore, if the interactions with the Ets factors CBF-A was originally described as a repressor of transcriptional activation. Furthermore, if the interactions with the Ets factors CBF-A was originally described as a repressor of transcriptional activation. Furthermore, if the interactions with the Ets factors.

In conclusion, CBF-A was purified due to its interaction with a positive control element within the SP6 transcriptional units. Whether the protein is only involved in transcriptional regulation or is most likely due to complex interactions with other regulatory elements in the different transcriptional units.

In conclusion, CBF-A was purified due to its interaction with a positive control element within the SP6 transcriptional units. Whether the protein is only involved in transcriptional regulation or is most likely due to complex interactions with other regulatory elements in the different transcriptional units. In conclusion, CBF-A was purified due to its interaction with a positive control element within the SP6 transcriptional units. Whether the protein is only involved in transcriptional regulation or is most likely due to complex interactions with other regulatory elements in the different transcriptional units. In conclusion, CBF-A was purified due to its interaction with a positive control element within the SP6 transcriptional units. Whether the protein is only involved in transcriptional regulation or is most likely due to complex interactions with other regulatory elements in the different transcriptional units. In conclusion, CBF-A was purified due to its interaction with a positive control element within the SP6 transcriptional units. Whether the protein is only involved in transcriptional regulation or is most likely due to complex interactions with other regulatory elements in the different transcriptional units. In conclusion, CBF-A was purified due to its interaction with a positive control element within the SP6 transcriptional units. Whether the protein is only involved in transcriptional regulation or is most likely due to complex interactions with other regulatory elements in the different transcriptional units. In conclusion, CBF-A was purified due to its interaction with a positive control element within the SP6 transcriptional units. Whether the protein is only involved in transcriptional regulation or is most likely due to complex interactions with other regulatory elements in the different transcriptional units. In conclusion, CBF-A was purified due to its interaction with a positive control element within the SP6 transcriptional units. Whether the protein is only involved in transcriptional regulation or is most likely due to complex interactions with other regulatory elements in the different transcriptional units. In conclusion, CBF-A was purified due to its interaction with a positive control element within the SP6 transcriptional units. Whether the protein is only involved in transcriptional regulation or is most likely due to complex interactions with other regulatory elements in the different transcriptional units. In conclusion, CBF-A was purified due to its interaction with a positive control element within the SP6 transcriptional units. Whether the protein is only involved in transcriptional regulation or is most likely due to complex interactions with other regulatory elements in the different transcriptional units. In conclusion, CBF-A was purified due to its interaction with a positive control element within the SP6 transcriptional units. Whether the protein is only involved in transcriptional regulation or is most likely due to complex interactions with other regulatory elements in the different transcriptional units. In conclusion, CBF-A was purified due to its interaction with a positive control element within the SP6 transcriptional units. Whether the protein is only involved in transcriptional regulation or is most likely due to complex interactions with other regulatory elements in the different transcriptional units. In conclusion, CBF-A was purified due to its interaction with a positive control element within the SP6 transcriptional units. Whether the protein is only involved in transcriptional regulation or is most likely due to complex interactions with other regulatory elements in the different transcriptional units. In conclusion, CBF-A was purified due to its interaction with a positive control element within the SP6 transcriptional units. Whether the protein is only involved in transcriptional regulation or is most likely due to complex interactions with other regulatory elements in the different transcriptional units. In conclusion, CBF-A was purified due to its interaction with a positive control element within the SP6 transcriptional units. Whether the protein is only involved in transcriptional regulation or is most likely due to complex interactions with other regulatory elements in the different transcriptional units.