MOLECULAR AND CELLULAR BIOLOGY, May 1991, p. 2665–2674
0270-7306/91/052665-10$02.00/0

Vol. 11, No. 5

Evi-1, a Murine Zinc Finger Proto-Oncogene, Encodes a Sequence-Specific DNA-Binding Protein

ARCHIBALD S. PERKINS,[1] RICHARD FISHEL,[2] NANCY A. JENKINS,[1] AND NEAL G. COPELAND[1]*

Mammalian Genetics Laboratory[1] and Molecular Mechanisms of Carcinogenesis Laboratory,[2] ABL-Basic Research Program, NCI-Frederick Cancer Research and Development Center, Frederick, Maryland 21702

Received 28 November 1990/Accepted 27 February 1991

***Evi-1* was originally identified as a common site of viral integration in murine myeloid tumors. *Evi-1* encodes a 120-kDa polypeptide containing 10 zinc finger motifs located in two domains 380 amino acids apart and an acidic domain located carboxy terminal to the second set of zinc fingers. These features suggest that Evi-1 is a site-specific DNA-binding protein involved in the regulation of RNA transcription. We have purified Evi-1 protein from *E. coli* and have employed a gel shift-polymerase chain reaction method using random oligonucleotides to identify a high-affinity binding site for Evi-1. The consensus sequence for this binding site is TGACAAGATAA. Evi-1 protein specifically protects this motif from DNase I digestion. By searching the nucleotide sequence data bases, we have found this binding site both in sequences 5′ to genes in putative or known regulatory regions and within intron sequences.**

Many eukaryotic transcriptional regulatory proteins contain one of several DNA-binding motifs, such as a helix-turn-helix or zinc fingers, which confer site-specific, high-affinity binding to transcriptional control sequences in genomic DNA. This binding, under proper conditions, is thought to result in changes in rates of transcriptional initiation by RNA polymerases by mechanisms that are not yet entirely clear (25, 37). The first two eukaryotic transcriptional regulatory proteins to be characterized, TFIIIA (17) and Sp1 (6), were identified by a combination of genetic and biochemical means, by which the binding sites were identified initially as *cis*-acting regulatory elements. Now that common DNA-binding motifs within proteins have been identified, putative transcriptional factors are often identified on the basis of protein sequence homology.

Evi-1 was first identified as a common site of ecotropic retrovirus integration in virally induced myeloid tumors of AKXD mice (40). Subsequent cloning and sequencing of the cDNA corresponding to this locus (39) showed that this locus encodes a 120-kDa polypeptide that contains 10 regions with extensive homology to the zinc-binding domains found in several transcriptional regulatory factors and first identified in the *Xenopus* transcription factor IIIA (36). This amino acid sequence consists of 27 to 30 residues with the consensus sequence (Y/F)$XCX_{2-4}CX_3FX_5LX_{2-3}HX_{3-4}HX_5$. Current structural models propose that the cysteine and histidine residues chelate a single zinc atom, constraining the intervening amino acids into a looped domain, which is further stabilized by hydrophobic interactions among (Y/F), F, and L (1, 30, 44). DNA binding then occurs between charged amino acids on the exposed face of an alpha helix (which extends from the center of the motif through the histidines) and an estimated 3 to 5 bases along the major groove (1, 7, 21). DNA-binding studies with purified zinc finger regions containing one or several fingers suggest that multiple fingers are necessary to obtain high-affinity, site-specific DNA binding, perhaps through cooperativity (18, 41, 44).

These sequence comparisons suggest that the *Evi-1* protein is a DNA-binding protein that acts to regulate mRNA transcription through interaction with one or more specific DNA sequence motifs. This notion is supported by the zinc-dependent nonspecific DNA-binding activity of Evi-1 reported recently (32). Although the exact role of *Evi-1* in myeloid leukemia is not clear, the effect of retroviral insertion is to dramatically increase the transcription of *Evi-1* in tumor cells, which suggests that increased levels of the protein product contribute to the transformed phenotype. The putative target DNA sequences for *Evi-1* high-affinity binding and action are not known.

In a previous report (45), we described results from Northern (RNA) and in situ hybridization analysis of *Evi-1* transcription in adult and embryonic mouse tissues. These data reveal a spatially and temporally restricted pattern of *Evi-1* expression which overlaps that of other putative transcription factors, such as *Hox-5.2*, *Hox-5.3* (14), and *Hox-7* (22, 49), as well as *Rarg* (retinoic acid receptor gamma) (15, 50). These findings suggest that *Evi-1* is an important regulatory molecule during morphogenesis.

As a first step towards identifying the target genes on which *Evi-1* may act, we sought to identify target sequences to which the Evi-1 protein binds. We have utilized a recently described technique (33) in which random sequence oligonucleotides are subjected to sequential cycles of gel shift and polymerase chain reaction (PCR) amplification to enrich for sequences which bind to the Evi-1 protein with high affinity. The sequences for several of these Evi-1-selected oligonucleotides were determined, defining a consensus recognition site. We show that bacterially expressed Evi-1 protein binds to this sequence in a site-specific manner.

MATERIALS AND METHODS

Bacteria, cells, plasmids, and oligonucleotides. *Escherichia coli* DH5 cells were purchased from Bethesda Research Laboratories. Bacterial strain MC1000 (8) was obtained from S. Brown. Plasmids pBluescript SK(−) and KS(+) were purchased from Stratagene. Plasmids pPR682 and pPR683 were kindly provided by P. Riggs (New England BioLabs). Plasmid pATH-2 was kindly provided by T. J. Koerner and A. Tzagaloff. Anti-TEV-1 antiserum was kindly provided by

* Corresponding author.

J. N. Ihle and T. Matsugi. Oligonucleotides were synthesized by Marilyn Powers on an Applied Biosciences model 380 B synthesizer.

Antipeptide antiserum and Western blot analysis. Amino acids 323 to 336 of Evi-1 (SP-317) were synthesized, purified, and injected into rabbits as described previously (10). High-titer antioligopeptide antiserum, as assessed by enzyme-linked immunosorbent assay, was obtained following the sixth boost. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gels were stained with 0.025% Coomassie blue R-250 or were electroblotted onto polyvinylidene difluoride membranes as described elsewhere (31). Following transfer, membranes were blocked with 0.25% gelatin in NET buffer (0.15 M NaCl, 5 mM EDTA, 50 mM Tris, pH 7.5, 0.2% Tween-20 [Sigma]) for 4 h, washed in NET buffer for 15 min, incubated with antiserum (1:250 dilution) for 4 h, washed in NET buffer for 15 min, incubated with 1 μCi of ^{125}I-protein A (Amersham) in NET buffer for 4 h, and washed for 45 min. Membranes were then dried briefly and autoradiographed.

Purification of MBP–Evi-1 protein. *Evi-1* was fused in frame to *MalE* of pPR683 (47a, 48) by first amplifying the coding region of *Evi-1* with primers 5′-CGGAATTCATGG CGCCTGACATC-3′ and 5′-CGTCTAGATTGGTCCCAC TCTG-3′ by PCR (51) with a full-length *Evi-1* cDNA as a template (p58.2-1 [39]). Twenty-five cycles, consisting of 1 min at 94°C, 2 min at 50°C, and 3 min at 72°C, with a 3-s extension of synthesis per cycle, were performed. This generated a 3.1-kb fragment of *Evi-1* spanning nucleotides 508 to 3636 (39) that encompassed the entire coding region. This fragment was purified, digested with *Eco*RI and *Xba*I, and ligated into *Eco*RI-*Xba*I-digested pPR683 (47a), which is identical to pMal-P (48) except for the polylinker region. Recombinant clones were identified by two criteria: (i) hybridization to a ^{32}P-labeled *Evi-1* probe and (ii) the presence of a novel 160-kDa band on Coomassie blue-stained SDS-PAGE analysis of whole-cell lysates of isopropyl-β-D-thiogalactopyranoside-induced bacteria that were immunoreactive with anti-Evi-1 antiserum. When this plasmid was grown to high density, the 160-kDa fusion protein was a small proportion of immunoreactive fusion proteins, and lower-molecular-weight species were found to accumulate. The fusion protein was purified from a 1-liter growth essentially as described by Riggs (48), except following affinity chromatography on an amylose resin, the proteins were further purified by dialysis against buffer A (25 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid [HEPES], pH 7.5, 1 mM dithiothreitol [DTT], 0.05 mM EDTA, 0.1 mM $ZnCl_2$, 50 mM KCl, and 10% glycerol) and then step fractionation with KCl on a 1-ml MonoQ (Pharmacia) column equilibrated with buffer A. Protein fractions were dialyzed against buffer A.

Purification of TrpE–Evi-1. We purified a TrpE–Evi-1 fusion protein from *E. coli* using two assays for Evi-1: (i) Western blot analysis using anti-Evi-1 antiserum (anti-SP317 and anti-TEV-1) and (ii) gel retardation assay using the Evi-1 binding consensus sequence as a substrate (see below). We constructed a *TrpE–Evi-1* fusion plasmid by inserting the entire coding sequence for *Evi-1* in frame downstream from *TrpE* in the plasmid pATH-2. This was achieved by amplifying the *Evi-1* coding region (nucleotides 460 to 3630 [39]) from p58.2-1 with primers 5′-CGGAATTCGATGAAGAGT GAAGAGGACCCG-3′ and 5′-GCATCGATCCACTCTGGT CAACCTTGACAA-3′ by PCR. DNA products were digested with *Eco*RI and *Cla*I, and a 3.2-kb product was purified. This fragment was first inserted into the polylinker of Bluescript SK(−) plasmid (Stratagene) in frame with the ATG within the Bluescript polylinker. We then transferred the *Evi-1* gene as a *Bam*HI-*Cla*I fragment to expression vector pATH-2, since the Bluescript construct failed to yield the expected unfused 120-kDa Evi-1 protein product when induced with isopropyl-β-D-thiogalactopyranoside. A plasmid clone containing *Evi-1* in frame with *TrpE*, designated pATH2-*Evi-1*, was selected, and analysis of whole-cell lysates of *E. coli* DH5 harboring this plasmid by SDS-PAGE revealed the presence of a novel 160-kDa band that was immunoreactive with anti-*Evi-1* antiserum (anti-SP317) on Western blot analysis.

For purification of TrpE–Evi-1 protein, a 2.5-liter culture of MC1000 strain *E. coli* containing pATH2-*Evi-1* was grown as described previously (13). Two to three hours postinduction with indole acrylic acid (10 μg/ml), the bacteria were lysed and sonicated as described by Riggs (48). This and all subsequent steps of the purification were conducted at 4°C. The crude lysate was clarified by centrifugation at 164,000 × g_{av}, and proteins were then precipitated by the addition of ammonium sulfate to 45% saturation and centrifugation at 12,000 × g_{av}. The pellet was resuspended in TDE buffer (25 mM Tris, pH 7.2, 1 mM DTT, 0.1 mM EDTA, and 10% glycerol) containing 0.1 M KCl, 0.1 mM $ZnCl_2$, and 0.25 mM phenylmethylsulfonyl fluoride, and was dialyzed for 4 to 5 h against this buffer. Triton X-100 was then added to a final concentration of 0.1%. The sample was then applied to a DE 52 column (2.5 by 30 cm) equilibrated with TDE containing 0.1 M KCl, 0.1% Triton X-100, and 0.25 mM phenylmethylsulfonyl fluoride. Proteins were eluted from the column at 24 ml/h with 0.25 and 0.5 M KCl steps. Active protein was present in both the 0.25 M KCl and 0.5 M KCl eluants, but since only the 0.25 M fraction was devoid of DNA contamination, it was selected for further purification.

The 0.25 M KCl fractions from the DE 52 column were pooled, dialyzed to 0.1 M KCl in TDET buffer (TDE buffer containing 0.1% Triton X-100), and applied to a phosphocellulose column (2.5 by 30 cm) which was equilibrated with TDET buffer plus 0.1 M KCl and run at 24 ml/h. The column was washed with an additional 2 to 3 volumes of equilibration buffer prior to TrpE–Evi-1 elution with a 0.6 M KCl step. Peak fractions were pooled and dialyzed against TDET buffer plus 300 mM KCl. Following dialysis, the 0.6 M KCl pooled phosphocellulose fractions were applied to a heparin-agarose (Sigma) column (1 by 10 cm) equilibrated in TDET with 0.3 M KCl and run at 12 ml/h. The column was washed with an additional 10 volumes of equilibration buffer, and the bound proteins were eluted with a 50-ml linear 0.3 to 1.0 M KCl gradient. TrpE–Evi-1 was eluted in fractions ranging from 370 to 405 mM KCl. Active fractions were pooled, dialyzed against TDET buffer plus 100 mM KCl, and applied to a MonoQ column (0.5 by 4 cm) equilibrated with buffer B (25 mM HEPES, pH 7.5, 1 mM DTT, 0.1 mM EDTA, 10% glycerol) plus 100 mM KCl and run at 12 ml/h. The column was washed with an additional 10 volumes of buffer B–100 mM KCl, and proteins were step eluted off the column with 0.1 M KCl increments. TrpE–Evi-1 was found exclusively in the 0.4 M KCl step. By Coomassie blue-stained SDS-PAGE, we estimate that this fraction is 95% pure.

Gel retardation assay and PCR cycles. The random and nonrandom oligonucleotides used as substrates for gel retardation were exactly as described by Mavrothalassitis et al. (33). Briefly, a 50-base single-strand synthetic oligonucleotide with the sequence 5′-CTGGATCCTAGATATCCCTG NNNNNNNNNNAGGCTCAAAGCTCAATTCCT-3′ was used as a template for second-strand synthesis, which was

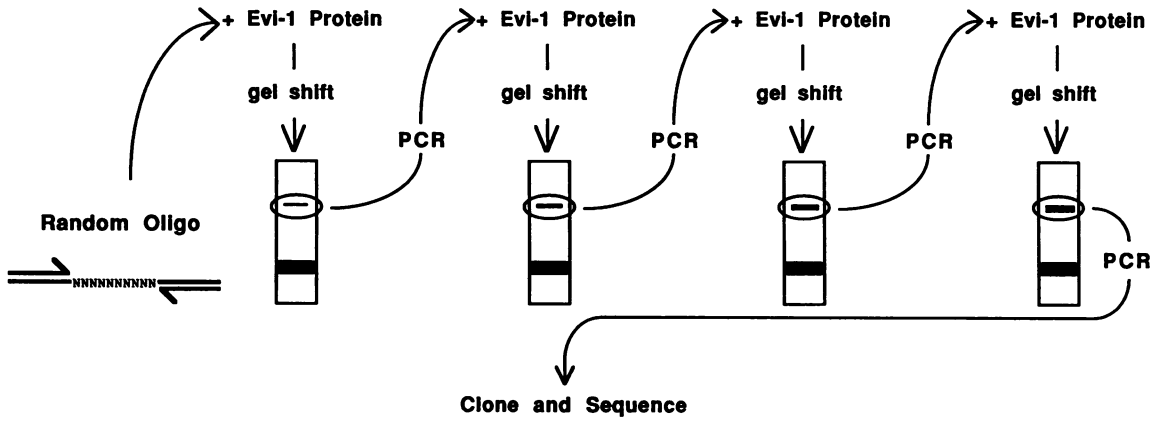

FIG. 1. Scheme for enrichment of high-affinity binding sites for Evi-1. Based on the technique of Mavrothalassitis et al. (33), this scheme entails four sequential cycles of gel shift and PCR that start with a 50-mer double-stranded oligonucleotide in which the central 10 nucleotides are random. Purified Evi-1 protein is allowed to bind to ^{32}P-labeled oligonucleotide, and then gel electrophoresis is performed to separate the bound from the unbound DNA. The shifted DNA is purified and amplified by PCR with primers that are homologous to the nonrandom sequences flanking the central random core. The PCR product is purified and subjected to another round of gel shift and PCR, etc. The final shifted band is again amplified and purified. This DNA is then cloned into a plasmid cloning vector and sequenced.

performed with a radiolabeled 20-mer (5'-AGGAATTCAG CTTTGAGCCT-3') as a primer and 5 U of Klenow fragment of DNA polymerase in the presence of 0.125 mM deoxyribonucleotide triphosphates. The resultant radiolabeled 50-bp DNA was purified by electrophoresis on 8% nondenaturing polyacrylamide, followed by elution and precipitation as described elsewhere (33). For the cycles of enrichment for high-affinity Evi-1 binding sites, the binding-reaction mixtures (10 μl) contained 25 mM HEPES, pH 7.5, 50 mM KCl, 4 mM $MgCl_2$, 1 mM DTT, 20% glycerol, 250 μg of bovine serum albumin (BSA) per ml, 250 μg of poly(dI-dC) per ml, 0.5×10^6 to 3×10^6 cpm of ^{32}P-labeled double-stranded oligonucleotide (approximately 3 to 15 ng), and 50 ng of protein. Gel retardation was done as described by Chodosh (9). The shifted band, representing the protein-DNA complex, was excised from the gel and eluted from the acrylamide by agitating overnight at 37°C with 0.6 ml of a solution containing 0.5 M ammonium acetate, 1 mM EDTA, 0.1% SDS, and 10 mM Tris-HCl, pH 7.5. DNA was isolated from the buffer by extraction with phenol-chloroform and precipitation with ethanol in the presence of 10 μg of yeast RNA. The DNA pellet was air-dried and resuspended in 10 μl of 10 mM Tris (pH 8)–1 mM EDTA (TE). PCRs (50 μl) were then performed under conditions suggested by the supplier (Perkin-Elmer-Cetus) and with 8 μl of the DNA recovered from gel retardation. Following the fourth cycle of gel shift and PCR, the resultant pool of 50-mers was digested with *Eco*RI and *Bam*HI, purified on an 8% polyacrylamide gel as described above, ligated into *Eco*RI-*Bam*HI-digested plasmid vector pBluescript KS+, and transformed into competent DH5α *E. coli* (Bethesda Research Laboratories) cells.

For the analytical gel retardation assay (see Fig. 3), the binding-reaction mixtures were as described above except that they contained 10,000 to 50,000 cpm of labeled oligonucleotide. Unlabeled double-stranded oligonucleotide for competition studies (see Fig. 7) was prepared by synthesizing a single-stranded oligonucleotide template with the sequence AGAGATCTAGCTTTGAGCCTGTTATCTTGTCA GGGATATCTAGGATCCAG. One nanomole was annealed with 1 nmol of the primer 5'-CTGGATCCTAGATATCCC TG-3' in 0.35 ml of 50 mM NaCl–10 mM Tris (pH 8.0)–0.1 mM EDTA and then adjusted to 10 mM $MgCl_2$ and 0.25 mM (each) deoxyribonucleotide triphosphates, and a second strand was synthesized with 100 U of Klenow fragment of DNA polymerase (Boehringer Mannheim). The products were separated on an 8% nondenaturing acrylamide gel, and the double-stranded 50-mer was eluted, extracted, precipitated, and quantitated by measuring the optical density at 260 nm.

DNase I footprinting. Bacterial extracts for DNase I footprinting were prepared as described by Landschulz et al. (29). An end-labeled DNA substrate was generated by first labeling the M13 reverse sequencing primer with ^{32}P and T4 polynucleotide kinase and then performing PCR on a pBluescript clone containing a selected oligonucleotide with the sequence TGACAAGATAA, with the labeled primer and an M13 forward sequencing primer. A radiolabeled 240-bp DNA product of the PCR was purified on low-melting-point agarose and DEAE chromatography and had a specific activity of 12×10^6 cpm/μg. DNase I footprinting was performed as described by Jones et al. (26). Binding-reaction mixtures (50 μl) contained 1 μg of poly(dI-dC), 20 μl of protein extract (approximately 100 μg of total protein), and 2,400 cpm of end-labeled probe.

RESULTS

Identification of an *Evi-1* binding site. The predicted amino acid sequence of Evi-1 suggests that it is a sequence-specific DNA-binding protein involved in the regulation of RNA transcription. To identify possible target sequences for Evi-1, we utilized a technique described by Oliphant et al. (43), as modified by Mavrothalassitis et al. (33), in which a 50-bp double-stranded oligonucleotide containing 20 bp of defined sequence on either side of a 10-bp stretch of random sequence is used as a ligand in gel retardation assays (Fig. 1). The bound DNA is then amplified by using the defined flanks as priming sites in a PCR. The amplified DNA is then subjected to three or four additional cycles of binding and amplification. The resultant DNA is then cloned, and several independent isolates are sequenced.

Evi-1 contains 10 zinc fingers separated into two domains; the amino-terminal domain (domain 1) contains 7 zinc fingers, while the more carboxy-terminal domain contains the

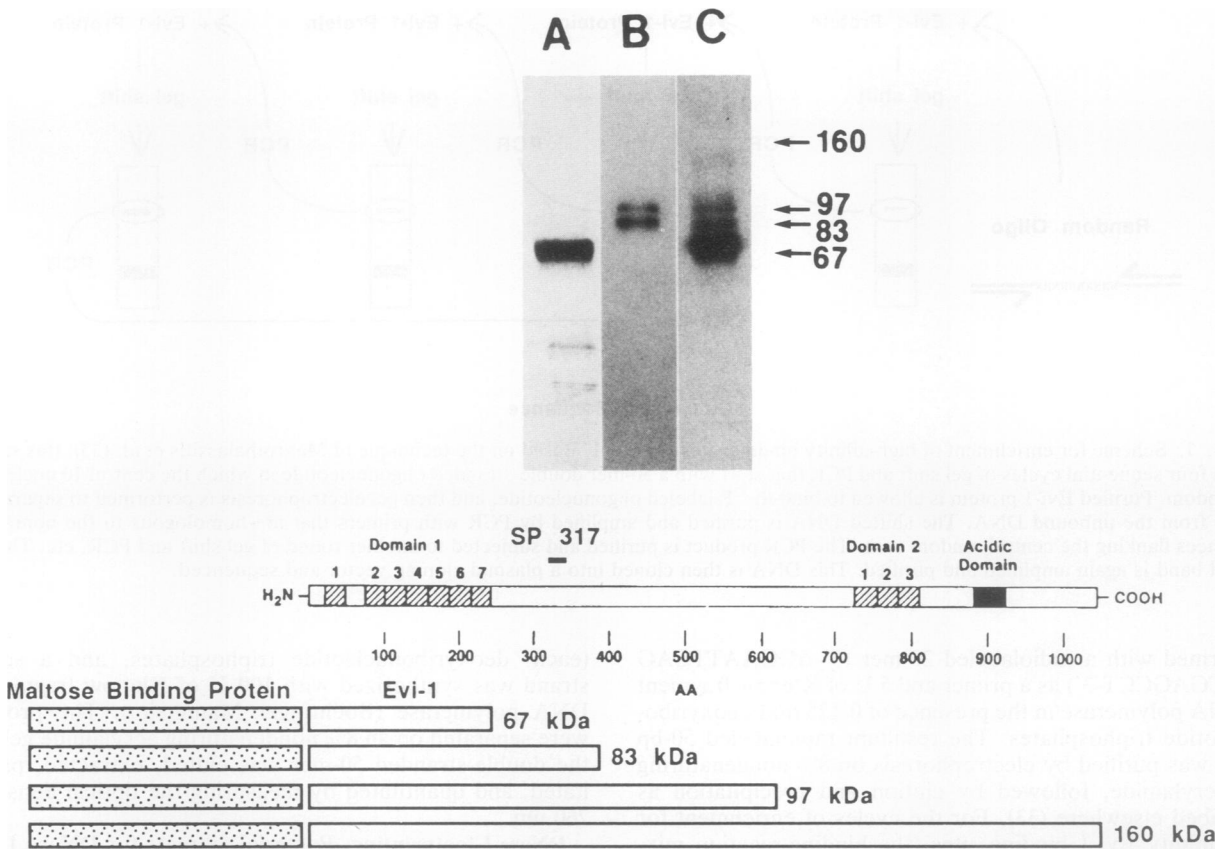

FIG. 2. SDS-PAGE and Western analysis of MBP–Evi-1. Lane A, Coomassie blue-stained SDS-PAGE; lane B, Western blot with anti-SP317 antiserum, which recognizes amino acids 323 to 336; lane C, Western blot with anti-TEV-1 antiserum, which is directed against amino acids 1 to 158. Relative molecular masses, in kilodaltons, are noted on the right. The lower panel depicts the structure of the *Evi-1*-encoded polypeptide and the predicted limits of the MBP–Evi-1 polypeptides. The hatched boxes represent the zinc fingers, which are numbered 1 to 7 in domain 1 and 1 to 3 in domain 2. The stippled box represents the *MalE*-encoded maltose-binding moiety, which constitutes 40 kDa of the fusion proteins. The adjoined open boxes depict the *Evi-1*-encoded portions of the three observed species (97, 83, and 67 kDa) and the predicted 160-kDa species. The major 67-kDa species terminates before amino acid 328, since it is not recognized by anti-SP317 antiserum. The 160-kDa species, although observed in bacterial preparations grown to lower density, was not observed in the preparation used for these studies.

remaining 3. A role for the zinc finger motifs of Evi-1 in DNA binding is suggested by zinc-dependent, nonspecific DNA-binding activity that has been described elsewhere (32). Since most sequence-specific DNA-binding zinc finger proteins contain fewer than 10 zinc fingers that are usually clustered in a single tandem array (21), it is likely that the two zinc finger clusters in Evi-1 bind separate sequences. With this possibility in mind, we used an Evi-1 protein preparation that primarily contained the first domain of Evi-1 to select for Evi-1 binding sequences. Such a protein preparation was obtained fortuitously as a fragment of a maltose-binding protein (MBP)–Evi-1 fusion construct.

A chimeric *MalE–Evi-1* fusion gene was created by inserting the full-length coding sequences of *Evi-1* into pPR683 (47a, 48). As a result of the *MalE*-encoded maltose-binding protein moiety at the amino terminus, the encoded MBP–Evi-1 fusion protein bound to an amylose affinity column (48). From a 1-liter culture, 5 mg of protein bound to the amylose resin, of which less than 10% was immunoreactive with anti-Evi-1 antisera while over 80% was 40 to 45 kDa in size and was not immunoreactive with anti-Evi-1 antisera. These smaller species were eluted from a MonoQ column with 0.1 to 0.3 M KCl, while Evi-1-immunoreactive protein was concentrated in the 1.0 M KCl fraction and migrated primarily as a 67-kDa protein on Coomassie blue-stained SDS gels (Fig. 2). This protein was immunoreactive only with antiserum TEV-1, which was generated against amino acids 1 to 158, and not with antiserum raised against amino acids 323 to 336 (anti-SP317; Fig. 2, lanes B and C). This pattern of immunoreactivity, together with the fact that the Evi-1 fusion protein binds to amylose, indicates that the protein is truncated and most likely contains only zinc finger domain 1 but not domain 2 (Fig. 2). However, we cannot rule out that this fraction contains a minute amount of protein, undetectable by Western blot and Coomassie blue-stained gels, that has domain 2. Truncation of MBP–Evi-1 was most likely due to either translational arrest or degradation of the protein. Degradation of foreign proteins fused to MBP has been observed previously (47b) and may be due to the presence of the secretion signal sequence, which directs the protein to the periplasmic space (19).

We used the 1.0 M KCl fraction of the MonoQ column, containing truncated MBP–Evi-1 protein, for the cycles of gel retardation and PCR. To control for nonspecific DNA binding, an end-labeled, double-stranded 50-mer oligonucleotide was synthesized with a defined (nonrandom) segment

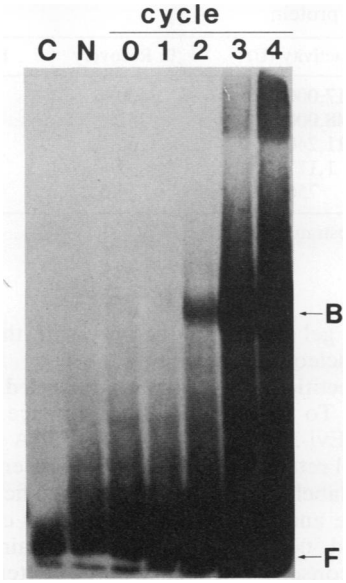

FIG. 3. Gel retardation assay on oligonucleotides from gel shift-PCR cycles. Either 0 ng (lane C) or 50 ng (lanes N and 0 to 4) of 67-kDa MBP–Evi-1 protein and 50,000 cpm of ^{32}P-labeled oligonucleotide were allowed to bind and were then fractionated on nondenaturing PAGE to separate bound (B) from free (F) oligonucleotides. In lane N, the substrate was a nonrandom oligonucleotide. In lanes 0 to 4, the substrate was random oligonucleotide enriched by 0 to 4 cycles of gel shift and PCR as indicated. In the absence of protein, all oligonucleotides migrated to position F, indistinguishable from the pattern in lane C, which contains cycle 4 oligonucleotide without protein. Although a retarded band is barely visible in lane 0 (random oligonucleotide), the intensity of the retarded band was more intense when preparative retardation assays were done, since 20 to 40 times as much radioactive ligand was used (data not shown).

of 10 nucleotides in the center. Binding-reaction conditions were optimized to minimize binding to the nonrandom oligonucleotide while still giving detectable levels of binding to the random oligonucleotide. This was done by increasing poly(dI-dC) in the binding-reaction mixture to 0.25 mg/ml and using 0.25 mg of BSA per ml. We found that magnesium chloride (0.5 to 4 mM) was essential for the binding reaction (data not shown). Binding-reaction mixtures also contained 50 mM KCl, 25 mM HEPES (pH 7.5), 1 mM DTT, and 20% glycerol. We conducted two additional controls to support our contention that the *Evi-1*-encoded moiety of the fusion protein in the 1.0 M KCl fraction is responsible for the gel shift activity. First, proteins purified from *E. coli* harboring pPR682 (which is identical to pPR683 except for the polylinker region) without *Evi-1* by amylose affinity chromatography and MonoQ chromatography did not show any gel shift activity. In addition, using anti-*Evi-1* antiserum (anti-TEV-1 [32]) we could retard the protein-DNA complex further in the gel retardation assay, while nonimmune control serum had no effect (data not shown).

After four cycles of gel retardation and amplification, we assessed the efficacy of enrichment for high-affinity *Evi-1* binding sites by performing gel retardation on DNA from each cycle. In each reaction, identical specific activities (3,000 Ci/mmol) and counts per minute (50,000 cpm) were present. We found a dramatic increase in the percentage of DNA shifted with each cycle (Fig. 3), strongly suggesting that we were successful in enriching for high-affinity sites.

The shifted DNA from the fourth cycle was inserted into a plasmid vector, and 16 independent clones were sequenced. There was remarkable identity found among the clones, and the consensus sequence TGACAAGATAA was easily discernable. Although the original pool of random oligonucleotides contained only 10 random nucleotides, an 11-bp consensus was obtained because the phasing and orientation of the sequence within the central 10 nucleotides (denoted in Fig. 4 by the boldface print) with respect to the nonrandom

RANDOM SUBSTRATE: A) ORIENTATION...GAGCCTNNNNNNNNNNCAGGGA...

B) ORIENTATION...TCCCTGNNNNNNNNNNAGGCTC...

SELECTED SEQUENCE:																ORIENTATION:
1	C	T	G	A	C	A	A	G	A	T	A	A	C	A	G	A
2	C	T	G	A	C	A	A	G	A	T	A	A	C	A	G	A
3	C	T	G	A	C	A	A	G	A	T	A	A	C	A	G	A
4	C	T	G	A	C	A	A	G	A	T	A	G	C	A	G	A
5	G	T	G	A	C	A	A	G	A	T	A	A	G	G	C	B
6	G	T	G	A	C	A	A	G	A	T	A	A	G	G	C	B
7	T	G	G	A	C	A	A	G	A	T	A	A	A	G	G	B
8	C	T	G	A	C	A	A	G	A	T	A	G	C	A	G	B
9	C	T	G	A	C	A	A	G	A	T	T	A	A	A	G	B
10	C	T	G	A	C	A	A	G	A	T	A	A	T	A	G	B
11	C	T	G	A	C	A	A	G	A	T	A	A	A	A	G	B
12	C	T	G	A	C	A	A	G	A	T	A	C	T	A	G	B
13	C	T	G	A	C	A	A	G	A	T	A	A	C	A	G	B
14	C	T	G	A	C	A	A	G	A	T	A	G	T	A	G	B
15	C	T	G	A	C	T	A	G	A	T	A	A	C	A	G	B
16	C	T	G	A	C	A	A	G	A	T	A	A	T	A	G	B
CONSENSUS:		T	G	A	C	A	A	G	A	T	A	A				
FREQUENCY:		15/16	16/16	16/16	16/16	15/16	16/16	16/16	16/16	16/16	15/16	12/16				

FIG. 4. Sequence analysis of 16 random plasmid clones containing oligonucleotides from the enriched pool. At the top is the sequence of the 6 nucleotides 5′ and 3′ to the central core of 10 random nucleotides, depicted in both orientations (A and B). Below this are shown the sequences of the central nucleotides within the oligonucleotide inserts of 16 plasmid clones. The limits of the central 10 bases are indicated by boldface letters, and the orientations of the inserts relative to the nonrandom flanking sequences are indicated on the right. At the bottom is the consensus sequence.

TABLE 1. Purification of TrpE–Evi-1 fusion protein

Fraction	Vol (ml)	Amt of protein (mg)	Sp act (U/mg)[a]	Total activity (U)[a]	% Recovery	Enrichment (fold)
Ammonium sulfate	20	534	31	17,000	100	1
DE 52, 0.25 M	21	223	214	48,000	282	6.9
Phosphocellulose	21	7	1,600	11,240	66	50
Heparin-agarose	4.6	0.33	3,500	1,177	7	113
MonoQ	3	0.1	7,100	756	4.5	230

[a] One unit is defined as the activity that will shift 1 fmol of probe in the presence of 16 fmol of substrate.

flanks was such that it yielded a very similar 11-bp sequence in all clones examined. The "A" orientation (Fig. 4) puts a T as the base adjacent to the central 10 nucleotides, and in all four clones with this orientation the central 10 nucleotides begin with G, so that the consensus is maintained. In the "B" orientation, in which the central 10 nucleotides were in the reverse orientation compared with the A orientation, there is a TG adjacent to the central 10 nucleotides, and in two clones, T was the first base (clones 5 and 6), while in nine clones, A was the first base (clones 8 to 16) of the central 10 nucleotides. In both of these cases, the consensus is maintained. Seven of the eleven positions in the consensus sequence were invariant. Positions 1, 5, and 10 each had one variation, while position 11 was more variant, with four variations.

Full-length Evi-1 binds specifically to the TGACAAGATAA site. We wished to test whether full-length Evi-1 protein also bound with high affinity to the TGACAAGATAA consensus sequence, and since we were unable to obtain full-length Evi-1 protein as a fusion with the maltose-binding protein, we utilized a *TrpE* fusion construct which in our hands yielded more full-length *Evi-1* fusion protein. The complete coding sequence for *Evi-1* was fused in frame to *TrpE* in plasmid pATH-2 to generate pATH-*Evi-1*. Two and one-half liters of *E. coli* harboring this plasmid was induced with indole acrylic acid, lysed with lysozyme and detergent, sonicated, and clarified by centrifugation to yield a cleared lysate fraction. Purification of the TrpE–Evi-1 fusion protein from this fraction involved an ammonium sulfate precipitation and four sequential chromatography steps (Table 1).

We assessed activity at each step of the purification by gel retardation assay using an end-labeled double-stranded 50-mer oligonucleotide containing the TGACAAGATAA consensus sequence, and concomitantly, we assessed the integrity and immunoreactivity of the protein to anti-Evi-1 antiserum by SDS-PAGE and Western analysis (Fig. 5). The heparin-agarose column was developed with a linear gradient of KCl (Fig. 5b), and as shown in Fig. 5, the gel retardation activity coeluted with a 160-kDa polypeptide immunoreactive with anti-Evi-1 antiserum, suggesting that the activity resided in the TrpE–Evi-1 fusion protein. In total, purification generated a 230-fold enrichment of DNA-binding activity. Analysis of equal numbers of gel shift units from each step in the purification by Coomassie blue-stained SDS-PAGE is shown in Fig. 6. The fairly identical intensities of the 160-kDa bands in each step indicate that there is no significant loss of specific activity during the purification. The final fraction (Fig. 6, lane 5) contained primarily the 160-kDa protein with minor contamination by a 44-kDa protein, which by two criteria is most likely an *E. coli*-encoded protein: (i) it was not immunoreactive with any of our anti-Evi-1 antisera, and (ii) a control protein extract from *E. coli* harboring the vector plasmid pATH-2 showed the presence of the 44-kDa band but not the 160-kDa band and did not exhibit gel retardation activity with the Evi-1 binding-site oligonucleotide.

Specific competition of binding with unlabeled TGACAAGA TAA sequence. To further show the sequence specificity of the binding of Evi-1 to the TGACAAGATAA sequence, we performed a gel retardation assay in the presence of various amounts of unlabeled specific and nonspecific competitors. Both the probe and the specific competitor consisted of a double-stranded 50-mer oligonucleotide containing the TGACAAGATAA consensus sequence at the center; the nonspecific competitor was sonicated mouse DNA (20 to 500 bp in length). Ten nanograms of MBP-Evi-1 protein was mixed with 120 fmol (7,000 cpm) of ^{32}P-labeled oligonucleotide probe in the presence of a 1-, 5-, 20-, or 100-fold excess of competitor. The presence of a fivefold excess of the specific competitor reduced the amount of bound, radiolabeled oligonucleotide by 80% relative to that in the binding reaction carried out without competitor (Fig. 7). By contrast, there was no decrease in the amount of bound oligonucleotide probe in the presence of a fivefold excess of the nonspecific competitor. Further decreases in the amount of bound probe were obtained in the presence of 20- and 100-fold excesses of the specific competitor.

DNase I protection of TGACAAGATAA sequence. Further evidence for specific binding of *Evi-1* to the TGACAAGA TAA consensus site was obtained by DNase I footprint analysis. A 240-bp double-stranded DNA fragment containing the consensus binding site was end labeled with ^{32}P on one strand. We prepared extracts from *E. coli* strains harboring pATH2, pATH2-*Evi-1*, pPR682, or pMBP-*Evi-1* and used these extracts to perform DNase I protection assays. Only the pATH2-*Evi-1* and pMBP-*Evi-1* extracts showed DNase I protection activity (Fig. 8, lanes 3 and 5), and as expected, the region pairs protected (from bp 124 to 145, relative to the labeled end) completely covers the TGACAA GATAA sequence (located from bp 129 to 140) and extends 5 bp to each side of it.

DISCUSSION

The identification of high-affinity binding sites in DNA for putative transcription factors is an important step in their characterization. In this paper, we present evidence that *Evi-1* encodes a sequence-specific DNA-binding protein that recognizes the sequence TGACAAGATAA. This binding activity most likely resides in the first domain of zinc fingers which contains 7 of the 10 zinc fingers. To identify the Evi-1 binding site, we have used a technique described recently by Mavrothalassitis et al. (33) which entails sequential gel retardation and PCR to enrich for high-affinity binding sites from a pool of approximately 10^6 random oligonucleotides. This approach is similar to, yet much simpler than, the technique described by Kinzler and Vogelstein (27), in which the substrate for DNA binding is sonicated genomic DNA

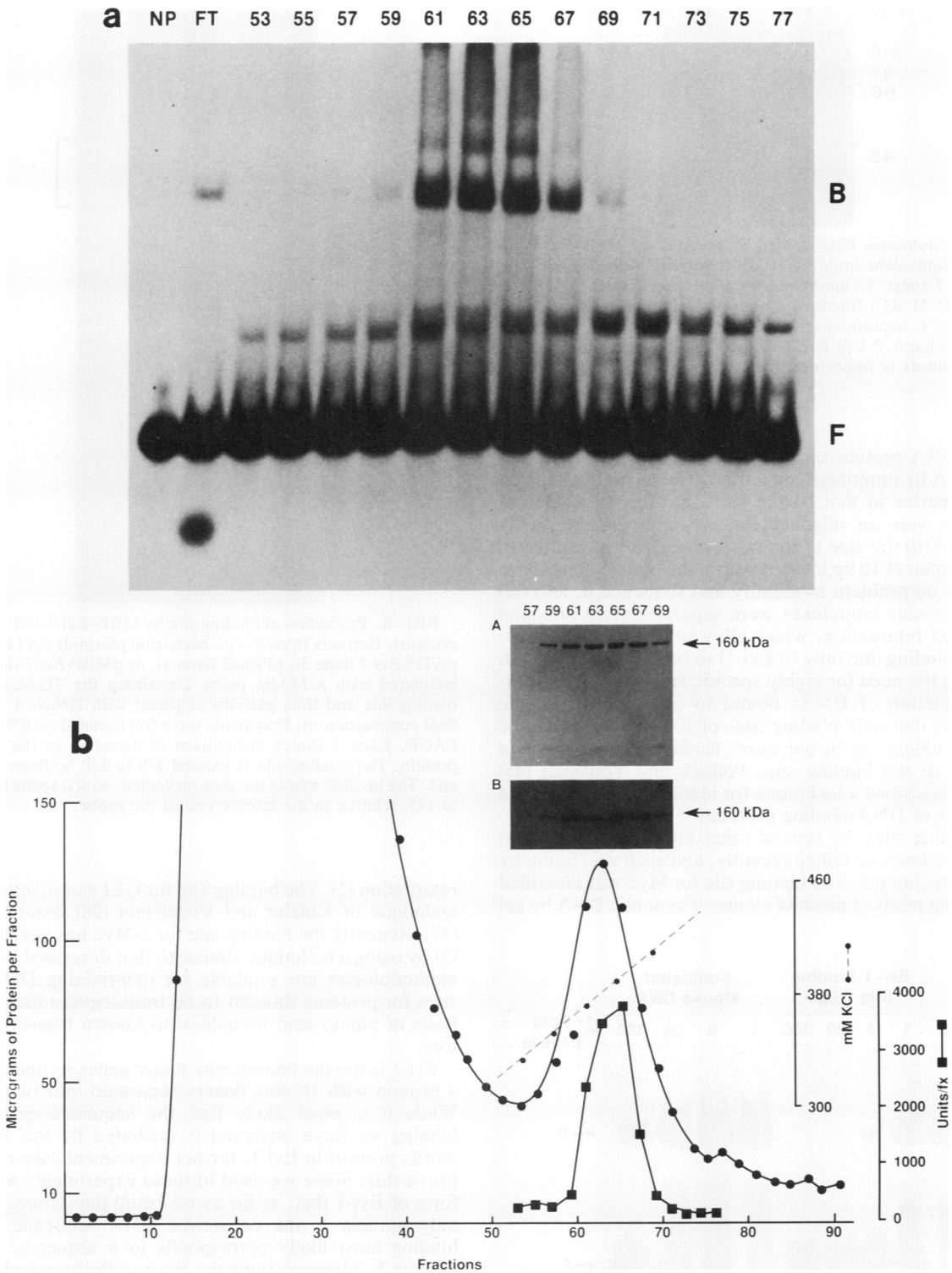

FIG. 5. (a) Gel retardation assay of fractions from heparin-agarose eluant. One microliter (0 to 50 ng) of eluant was allowed to bind to 10,000 cpm of ^{32}P-labeled double-stranded oligonucleotide containing the consensus binding sequence and then analyzed by nondenaturing PAGE to separate Evi-1-bound (B) from free (F) probes. The fraction number analyzed is indicated above each lane; NP designates a no-protein control; FT designates a flowthrough fraction. The protein(s) present in the band of intermediate mobility is an *E. coli*-encoded contaminant which is purified away in the last step of the purification. (b) Elution profile of heparin-agarose column chromatography. Units/fx, units of gel shift activity (as defined in Table 1) per fraction. (Insert A) Coomassie blue-stained polyvinylidine difluoride membrane to which proteins from an SDS-PAGE gel were electroblotted; (insert B) Western blot analysis of the same filter, probed with anti-SP317 antiserum. The fractions analyzed are indicated at the top of each lane.

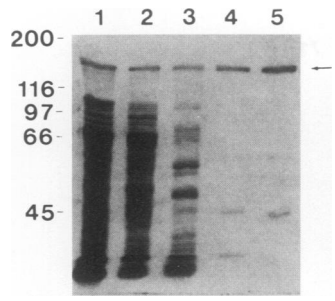

FIG. 6. Coomassie blue-stained SDS-PAGE of TrpE–Evi-1 purification. Equivalent units of gel shift activity were loaded onto each lane. Lanes: 1, ammonium sulfate precipitate; 2, DE 52 column, 0.25 M KCl fraction; 3, phosphocellulose column, 0.6 M KCl fraction; 4, heparin-agarose column, fractions 61 to 67 pooled; 5, MonoQ column, 0.4 M KCl fraction. The migration of molecular weight standards is indicated to the left, and the arrow indicates TrpE–Evi-1.

and the DNA-protein complexes are separated from unbound DNA by immunoprecipitation. The technique utilized here is superior in that (i) the substrate for the cycles of purification was an oligonucleotide which can be easily synthesized (ii) the size of the random region was limited to a discrete span of 10 bp located within the oligonucleotide so that it was no problem to identify and sequence it, and (iii) the DNA-protein complexes were separated from unbound DNA by gel retardation, which allowed a single band (representing binding due only to Evi-1) to be excised, thus both eliminating the need for highly specific antiserum and avoiding the isolation of DNAs bound to other proteins. One drawback is that only binding sites of 10 bp or fewer can be identified, unless, as in our case, flanking nucleotides can contribute to the binding site. Pollock and Treisman (46) have also described a technique for identifying the sequence specificities of DNA-binding proteins.

The binding sites for several other putative transcription factors have been identified recently, and each was found by a different technique. The binding site for Myb was identified by screening pools of plasmid clones of genomic DNA by gel

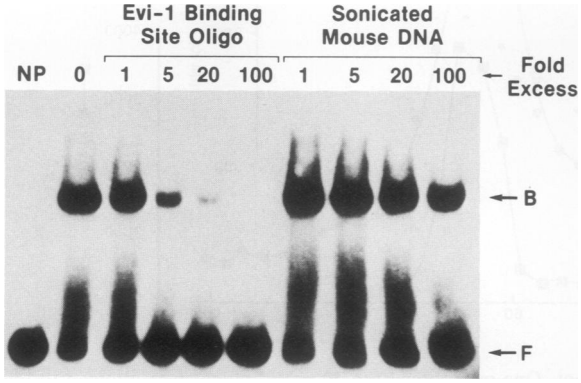

FIG. 7. Gel retardation assay in the presence of competitor DNAs. Purified MBP–Evi-1 fusion protein was incubated with radiolabeled 50-mer oligonucleotide containing the sequence TGACAAGATAA in the presence or absence of specific (Evi-1 binding site oligo) and nonspecific (sonicated mouse DNA) DNA competitors. Fold excess is relative to the labeled probe. In lane NP, no protein was included in the reaction mixture. B and F denote bound and free DNA, respectively.

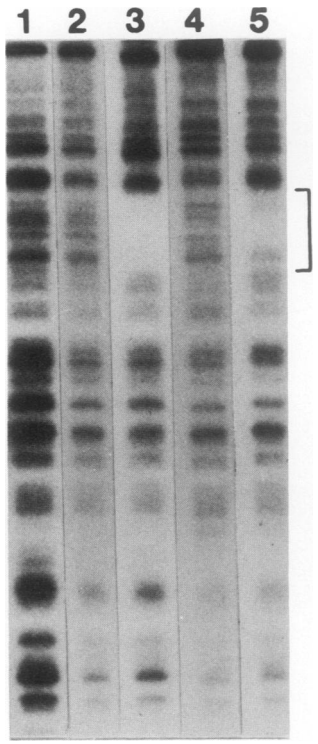

FIG. 8. Protection of binding site by MBP–Evi-1 and TrpE–Evi-1 proteins. Extracts from *E. coli* harboring plasmids pATH-2 (lane 2), pATH-*Evi-1* (lane 3), pPR682 (lane 4), or pMBP-*Evi-1* (lane 5) were incubated with a 240-bp probe containing the TGACAAGATAA binding site and then partially digested with DNase I (100 ng/ml, final concentration). Fragments were fractionated on 8% denaturing PAGE. Lane 1 shows the pattern of digestion in the absence of protein. The binding site is located 129 to 140 bp from the labeled end. The bracket marks the area protected, which spans from bp 124 to 145, relative to the labeled end of the probe.

retardation (2). The binding site for GLI was identified by the technique of Kinzler and Vogelstein (28) described above (27). Recently the binding site for c-Myc has been described (3) by using a technique similar to that described here. Thus, methodologies are available for determining DNA-binding sites for proteins thought to be transcription factors on the basis of amino acid homology to known transcription factors.

Evi-1 is unique among zinc finger genes in that it encodes a protein with 10 zinc fingers separated into two domains. While it is most likely that the sequence-specific DNA binding we have observed is mediated by the zinc finger motifs present in Evi-1, further experiments are required to prove this. Since we used in these experiments a truncated form of Evi-1 that, as far as we could determine, contained only domain 1, the consensus DNA sequence for Evi-1 binding most likely corresponds to a sequence bound by domain 1. Although unlikely, the possibility remains that an undetectable amount of domain 2-containing protein was present in the MBP–Evi-1 protein preparation that was used to identify the binding site and that the sequence found corresponds to a binding site for domain 2. To address this possibility more directly, we are creating maltose binding protein fusions with different regions of Evi-1 and will test the ability of these proteins to bind to the sequence we have identified. Since it is estimated that each zinc finger interacts

with 3 to 5 bases (21, 42), one would predict that only two to four fingers bind to the 11-bp binding site we have identified. We are currently investigating which of the seven zinc fingers in domain 1 are responsible for the site-specific binding we have observed. It is likely that *Evi-1* may bind with higher affinity to a site larger than 11 bp and that these experiments have revealed only the "core" to which *Evi-1* binds with highest affinity. To determine whether sequences flanking the 11 bp are important for binding, we are repeating the gel retardation-PCR cycles with oligonucleotides in which the 11-bp binding site is flanked by random nucleotides. Such an approach was used recently by Blackwell and Weintraub (4) to further define the binding sites for MyoD and E2A.

It remains possible that domain 2, which contains three zinc fingers, binds with high affinity to an unrelated DNA sequence. We are currently investigating this possibility. Several splicing patterns of *Evi-1* have been reported (5, 38). One of these splicing patterns results in an mRNA that has the potential to encode a 718-amino-acid polypeptide that lacks amino acids 190 to 513 of the full-length Evi-1 encoded by the 4.8-kb cDNA reported by Morishita et al. (39). The deleted region includes zinc fingers 6 and 7 of domain 1, and thus, the protein would have 8 instead of 10 zinc fingers. It is possible that this 718-amino-acid form of Evi-1 binds with altered affinity to the consensus sequence we have identified or that it binds to an unrelated sequence. In addition, the sequence from amino acids 258 to 713 of full-length Evi-1 is proline rich (58 prolines in 456 amino acids), and the 718-amino-acid form of Evi-1 lacks 28 of these prolines. Proline-rich regions of transcriptional regulatory proteins have been proposed to play a role in transcriptional activation (35). Thus, the alternatively spliced form of *Evi-1* may differ from the larger form both in its DNA-binding characteristics and in its effects on transcription.

Several transcriptional activators in yeast cells contain acidic domains that are required for inducing activation of transcription, most likely by interacting with other proteins in the transcriptional complex (reviewed in reference 47). The presence of an acidic domain adjacent to domain 2 suggests that Evi-1 can function in vivo as a transcriptional activator. We are currently investigating this possibility by constructing promoter sequences with the Evi-1 binding site linked to reporter genes and assaying for Evi-1 *trans*-activation by cotransfection.

To our knowledge, the DNA sequence we have identified as an Evi-1 binding site is distinct from the DNA-binding sites for other DNA-binding proteins (37). By searching the GenBank and EMBL DNA data bases using the software package of the Genetics Computer Group (12) for the sequence TGACAAGATAA or close variations thereof, we have found that this sequence or variations of it are present 5′ to several genes and within intron sequences of several others. For example, the human gene for proliferating cell nuclear antigen contains the site TCACAAGATAA at −400 nucleotides from the start of transcription (53). Proliferating cell nuclear antigen is a component of the cellular replication machinery and is found in cells actively undergoing division (reviewed in reference 52). The human gene for the alpha subunit for chorionic gonadotropin contains the sequence AAACAAGATAA at −155 nucleotides from the start of transcription, which is within the upstream regulatory element, and is situated adjacent to the cyclic AMP response elements (11, 24). Recently, Jameson et al. (23) have shown that two distinct proteins bind to the upstream regulatory element, and one was shown by methylation interference assay to contact the DNA over the AAACAAGATAA sequence.

This consensus sequence is also present within intron sequences of *Kras*, *Lmyc*, the erythropoietin gene (*Epo*), and the interleukin-6 gene. The sites of *Evi-1* expression (45) overlap with the sites of expression of some of these genes. For example, both *Evi-1* and *Epo* are expressed in the tubular cells in the kidney (34). In addition, both *Evi-1* (38) and human chorionic gonadotropin (16) have been found to be expressed in human endometrial carcinomas. We are currently investigating the significance of these findings. It is tempting to speculate that aberrant expression of *Evi-1* in leukemic cells leads to increased transcription of genes such as *PCNA*, *Kras*, *Lmyc*, *Il6*, or *Epo* and that increased levels of the products of these genes contribute to the transformed phenotype. Interestingly, enhanced expression of *Kras* due to Friend virus integration has been found in a murine early myeloid cell line and is thought to contribute to the transformed phenotype (20).

ACKNOWLEDGMENTS

This research was sponsored by the National Cancer Institute, DHHS, under contract no. NO1-CO-74101 with ABL.

We thank Peter F. Johnson for his valuable input and suggestions. We also thank Linda Brubaker for secretarial support, and we thank Arthur Buchberg, John Mercer, and Peter Johnson for critically reviewing the manuscript. We are indebted to George Mavrothalassitis for sharing his results prior to publication.

REFERENCES

1. **Berg, J. M.** 1988. Proposed structure for the zinc-binding domains from transcription factor IIIA and related proteins. Proc. Natl. Acad. Sci. USA **85**:99–102.
2. **Biedenkapp, H., U. Borgmeyer, A. E. Sippel, and K. M. Klempnauer.** 1988. Viral *myb* oncogene encodes a sequence-specific DNA-binding activity. Nature (London) **335**:835–837.
3. **Blackwell, T. K., L. Kretzner, E. M. Blackwood, R. N. Eisenman, and H. Weintraub.** 1990. Sequence-specific DNA-binding by the c-Myc protein. Science **250**:1149–1151.
4. **Blackwell, T. K., and H. Weintraub.** 1990. A new binding-site selection technique reveals differences and similarities between MyoD and E2A DNA-binding specificities. Science **250**:1104–1110.
5. **Bordereaux, D., S. Fichelson, P. Tambourin, and S. Gisselbrecht.** 1990. Alternative-splicing of the *Evi-1* zinc finger gene generates mRNAs which differ by the number of zinc finger motifs. Oncogene **5**:925–927.
6. **Briggs, M. R., J. T. Kadonoga, S. P. Bell, and R. Tjian.** 1986. Purification and biochemical characterization of the promoter-specific transcription factor, *Sp1*. Science **234**:47–52.
7. **Brown, R. S., and P. Argos.** 1986. Fingers and helices. Nature (London) **324**:215.
8. **Casadaban, M. J., and S. N. Cohen.** 1980. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J. Mol. Biol. **138**:179–207.
9. **Chodosh, L.** 1988. Mobility shift DNA-binding assay using gel electrophoresis, p. 12.2.1–12.2.7. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology. John Wiley & Sons, Inc., New York.
10. **Copeland, T. D., W.-P. Tsai, Y. D. Kim, and S. Oroszlan.** 1986. Envelope proteins of human T cell leukemia virus type I: characterization by antisera to synthetic peptides and identification of a natural epitope. J. Immunol. **137**:2945–2951.
11. **Delegeane, A. M., L. H. Ferland, and P. L. Mellon.** 1987. Tissue-specific enhancer of the human glycoprotein hormone α-subunit gene: dependence on cyclic AMP-inducible elements. Mol. Cell. Biol. **7**:3994–4002.
12. **Devereaux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic

Acids Res. 12:387–395.
13. **Dieckmann, C. L., and A. Tzagoloff.** 1985. Assembly of the mitochondrial membrane system *CBP6*, a yeast nuclear gene necessary for synthesis of cytochrome *b*. J. Biol. Chem. **266:** 1513–1520.
14. **Dollé, P., and D. Duboule.** 1989. Two gene members of the murine *Hox*-5 complex show regional and cell-type specific expression in developing limbs and gonads. EMBO J. **8:**1507–1515.
15. **Dollé, P., E. Ruberte, P. Kastner, M. Petkovich, C. M. Stoner, L. J. Gudas, and P. Chambon.** 1989. Differential expression of genes encoding α, β, and γ retinoic acid receptors and CRABP in the developing limbs of the mouse. Nature (London) **342:**702–705.
16. **Donaldson, E. S., J. R. van Nagell, Jr., S. Pursell, E. C. Gay, W. R. Mecker, R. Kashmiri, and J. Vande Voorde.** 1980. Multiple biochemical markers in patients with gynecologic malignancies. Cancer **45:**948.
17. **Engelke, D. R., S.-Y. Ng, B. S. Shastry, and R. G. Roeder.** 1980. Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. Cell **19:**717–728.
18. **Frankel, A. D., J. M. Berg, and C. O. Pabo.** 1987. Metal-dependent folding of a single zinc finger from transcription factor IIIA. Proc. Natl. Acad. Sci. USA **84:**4841–4845.
19. **Gentz, R., Y. Kuys, C. Zwieb, D. Taatjes, H. Taatjes, W. Bannwarth, D. Stueber, and I. Ibrahimi.** 1988. Association of degradation and secretion of three chimeric polypeptides in *Escherichia coli*. J. Bacteriol. **170:**2212–2220.
20. **George, D. L., B. Glick, S. Trusko, and N. Freeman.** 1986. Enhanced c-ki-*ras* expression associated with Friend virus integration in a bone marrow-derived mouse cell line. Proc. Natl. Acad. Sci. USA **83:**1651–1655.
21. **Gibson, T. J., J. P. M. Postma, R. S. Brown, and P. Argos.** 1988. A model for the tertiary structure of the 28 residue DNA-binding motif ('zinc-finger') common to many eukaryotic transcriptional regulatory proteins. Protein Eng. **2:**209–218.
22. **Hill, R. E., P. F. Jones, A. R. Rees, C. M. Sime, M. J. Justice, N. G. Copeland, N. A. Jenkins, E. Graham, and D. R. Davidson.** 1989. A new family of mouse homeobox-containing genes: molecular structure, chromosomal location, and developmental expression of *Hox*-7.1. Genes Dev. **3:**26–37.
23. **Jameson, J. L., C. Albanese, and J. F. Habener.** 1989. Distinct adjacent protein-binding domains in the glycoprotein hormone α gene interact independently with a cAMP-responsive enhancer. J. Biol. Chem. **264:**16190–16196.
24. **Jameson, J. L., R. C. Jaffe, P. J. Deutsch, C. Albanese, and J. F. Habener.** 1988. The gonadotropin α-gene contains multiple protein binding domains that interact to modulate basal and cAMP-responsive transcription. J. Biol. Chem. **263:**9879–9886.
25. **Johnson, P. F., and S. L. McKnight.** 1989. Eukaryotic transcriptional regulatory proteins. Annu. Rev. Biochem. **58:**799–839.
26. **Jones, K. A., K. R. Yamamoto, and R. Tjian.** 1985. Two distinct transcription factors bind to the HSV thymidine kinase promoter *in vitro*. Cell **42:**559–572.
27. **Kinzler, K. W., and B. Vogelstein.** 1989. Whole genome PCR: application to the identification of sequences bound by regulatory proteins. Nucleic Acids Res. **17:**3645–3653.
28. **Kinzler, K. W., and B. Vogelstein.** 1990. The *GLI* gene encodes a nuclear protein which binds specific sequences in the human genome. Mol. Cell. Biol. **10:**634–642.
29. **Landschulz, W. H., P. F. Johnson, and S. L. McKnight.** 1989. The DNA binding domain of the rat liver nuclear protein c/EBP is bipartite. Science **243:**1681–1688.
30. **Lee, M. S., G. P. Gippert, K. V. Soman, D. A. Case, and P. E. Wright.** 1989. Three-dimensional solution structure of a single zinc-finger DNA-binding domain. Science **245:**635–637.
31. **Matsudaira, P.** 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. J. Biol. Chem. **262:**10035–10038.
32. **Matsugi, T., K. Morishita, and J. N. Ihle.** 1990. Identification, nuclear localization, and DNA-binding activity of the zinc finger protein encoded by the *Evi-1* myeloid transforming gene. Mol. Cell. Biol. **10:**1259–1264.
33. **Mavrothalassitis, G., G. Beal, and T. S. Papas.** 1990. Defining target sequence(s) of a DNA binding protein by random selection and PCR: determination of the GCN4 binding sequence repertoire. DNA Cell Biol. **9:**783–788.
34. **Maxwell, A. P., T. R. Lappin, C. F. Johnston, J. M. Bridges, and M. G. McGeown.** 1990. Erythropoietin production in kidney tubular cells. Br. J. Haematol. **74:**535–539.
35. **Mermod, N., E. A. O'Neill, T. J. Kelly, and R. Tjian.** 1989. The proline-rich transcriptional activator of CTF/NF-1 is distinct from the replication and DNA binding domains. Cell **58:**741–753.
36. **Miller, J., A. D. McLachlan, and A. Klug.** 1985. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. EMBO J. **4:**1609–1614.
37. **Mitchell, P. J., and R. Tjian.** 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science **245:**371–378.
38. **Morishita, K., E. Parganas, E. C. Douglass, and J. N. Ihle.** 1990. Unique expression of the human *Evi-1* gene in an endometrial carcinoma cell line and sequence of cDNAs and structure of alternatively spliced transcripts. Oncogene **5:**963–972.
39. **Morishita, K., D. S. Parker, M. L. Mucenski, N. A. Jenkins, N. G. Copeland, and J. N. Ihle.** 1988. Retroviral activation of a novel gene encoding a zinc-finger protein in IL-3-dependent myeloid leukemia cell lines. Cell **54:**831–840.
40. **Mucenski, M. L., B. A. Taylor, J. N. Ihle, J. W. Hartley, H. C. Morse III, N. A. Jenkins, and N. G. Copeland.** 1988. Identification of a common ecotropic viral integration site, *Evi-1*, in the DNA of AKXD murine myeloid tumors. Mol. Cell. Biol. **8:**301–308.
41. **Nagai, K., Y. Nakaseko, K. Nasmyth, and D. Rhodes.** 1988. Zinc-finger motifs expressed in *E. coli* and folded *in vitro* direct specific binding to DNA. Nature (London) **332:**284.
42. **Nardelli, J., J. J. Gibson, C. Vesque, and P. Charnay.** 1991. Base sequence discrimination by zinc-finger DNA-binding domains. Nature (London) **349:**175–178.
43. **Oliphant, A. R., C. J. Brandl, and K. Struhl.** 1989. Defining the sequence specificity of DNA-binding proteins by selecting binding sites from random-sequence oligonucleotides: analysis of yeast GCN4 protein. Mol. Cell. Biol. **9:**2944–2949.
44. **Párraga, G., S. J. Horvath, A. Eisen, W. E. Taylor, L. Hood, E. T. Young, and R. E. Klevit.** 1988. Zinc-dependent structure of a single-finger domain of yeast ADR1. Science **241:**1489–1492.
45. **Perkins, A. S., J. A. Mercer, N. A. Jenkins, and N. G. Copeland.** 1991. Patterns of *Evi-1* expression in embryonic and adult tissues suggest that *Evi-1* plays an important regulatory role in mouse development. Development, in press.
46. **Pollock, R., and R. Treisman.** 1990. A sensitive method for the determination of protein-DNA binding specificities. Nucleic Acids Res. **18:**6197–6209.
47. **Ptashne, M., and A. A. Gann.** 1990. Activators and targets. Nature (London) **346:**329–331.
47a. **Riggs, P.** Unpublished data.
47b. **Riggs, P.** Personal communication.
48. **Riggs, P. D.** 1990. Expression and purification of maltose-binding protein fusions, p. 16.6.1–16.6.10. *In* F. M. Ausubel et al. (ed.), Current protocols in molecular biology. Greene Publishing Associates/Wiley Interscience, New York.
49. **Robert, B., D. Sassoon, B. Jacq, W. Gehring, and M. Buckingham.** 1989. *Hox-7*, a mouse homeobox gene with a novel pattern of expression during embryogenesis. EMBO J. **8:**91–100.
50. **Ruberte, E., P. Dolle, A. Zelent, G. Morriss-Kay, and P. Chambon.** 1990. Specific spatial and temporal distribution of retinoic and receptor gamma transcripts during mouse embryogenesis. Development **108:**213–222.
51. **Scharf, S. J., G. T. Horn, and H. A. Erlich.** 1986. Direct cloning and sequence analysis of enzymatically amplified genomic sequences. Science **233:**1076–1078.
52. **Stillman, B.** 1989. Initiation of eukaryotic DNA replication *in vitro*. Annu. Rev. Cell Biol. **5:**197–245.
53. **Travali, S., D.-H. Ku, M. G. Rizzo, L. Ottavio, R. Baserger, and B. Calabretti.** 1989. Structure of the human gene for proliferating cell nuclear antigen. J. Biol. Chem. **264:**7466–7472.