MOLECULAR AND CELLULAR BIOLOGY, Dec. 1992, p. 5632–5639
0270-7306/92/125632-08$02.00/0

Vol. 12, No. 12

Novel Member of the Zinc Finger Superfamily: A C_2-HC Finger That Recognizes a Glia-Specific Gene

JIN G. KIM AND LYNN D. HUDSON*

Laboratory of Viral and Molecular Pathogenesis, National Institute of Neurologic Disorders and Stroke, Bethesda, Maryland 20892

Received 30 June 1992/Returned for modification 19 August 1992/Accepted 21 September 1992

A novel member of the zinc finger superfamily was cloned by virtue of its binding to *cis*-regulatory elements of a glia-specific gene, the myelin proteolipid protein (PLP) gene. Named MyTI (myelin transcription factor I), this gene is most highly transcribed in the developing nervous system, where expression precedes induction of its presumptive target, PLP. Low levels of MyTI transcripts can be detected in nonneural tissues only by polymerase chain reaction analysis. Zinc is a necessary cofactor for DNA binding of MyTI, as the zinc-chelating agent 1,10-orthophenanthroline eliminates binding activity. Zinc may stabilize the DNA-binding domain of MyTI by coordinating three cysteine and one histidine residue in a Cys-X_5-Cys-X_{12}-His-X_4-Cys (C_2-HC) arrangement. The MyTI protein has six fingers of the C_2-HC class arranged in two widely separated clusters. These two domains of DNA binding can function independently and recognize the same DNA sequence, suggesting that MyTI may contribute to the higher-order structure of a target promoter by simultaneously binding both proximal and distal sites. The six fingers are highly conserved, suggesting that they arose from successive duplication events, while the linker regions diverge in size and sequence. Both amino acid sequence comparisons and secondary-structure predictions indicate that the C_2-HC fingers of MyTI do not resemble the zinc-mediated loops of C_2-H_2 fingers, C_2-C_2 fingers, or C_x clusters. MyTI may therefore be the prototype of a new structural family of zinc-stabilized DNA binding proteins.

The mechanisms that dictate the final program of gene expression in a fully differentiated cell can be revealed by starting at either end of the regulatory cascade. Mutations that disrupt segmentation or the pattern of adult structures have defined hierarchies of regulatory genes, some of which bind DNA and may function as transcriptional regulatory factors. The complementary approach begins with the targets of regulation, the highly specialized tissue-specific genes expressed in terminally differentiated cells, and works backwards, or "bottom up" (35), to clone the pertinent regulatory genes and subject them to reverse genetic functional assays. To examine the series of controls operating on progenitors of the glial lineage, we have begun with one of the final targets of regulation in myelinating glial cells, proteolipid protein (PLP). PLP is the most abundant protein of central nervous system myelin, a highly specialized extension of the oligodendrocyte plasma membrane that ensheaths axons to permit the rapid saltatory conduction of impulses. This highly conserved protein plays a role in the structure and compaction of the myelin sheath (reviewed in reference 11). The developmental stage- and tissue-specific expression of PLP is coordinately controlled with expression of the other myelin proteins synthesized by oligodendrocytes in the central nervous system; however, the molecular mechanisms of this regulated expression are unknown.

Sequence-specific interactions between DNA-binding proteins and *cis*-acting elements have been shown to be important determinants of gene expression (6, 23). To characterize protein-DNA interactions involved in the control of PLP expression, we previously analyzed the human PLP promoter biochemically by gel shift and footprint assays and functionally by transfecting glial cells with PLP-chloramphenicol acetyltransferase chimeric genes (2). Only a few hundred bases of the PLP upstream region are necessary for maximal PLP expression in transfected cells, and located within this region are five binding sites for putative *trans*-acting proteins (2). These sites are conserved between species, and they encompass regions of the rat PLP promoter that also bind nuclear proteins (27). Several of these sites are homologous to regions found in other myelin-specific genes and may form the basis for the coordinate control of these genes.

To isolate genes encoding transcription factors for the human PLP promoter, a phage expression library was screened by a method that relies on the detection of DNA-protein interactions (17, 36). We report here the isolation of a human cDNA clone encoding a novel zinc finger protein (myelin transcription factor I [MyTI]) that binds to the proximal region of the PLP promoter.

MATERIALS AND METHODS

Screening of cDNA libraries and sequencing of inserts. Five DNA sites that bind nuclear proteins in the PLP promoter (site 1 [+3 to −20], site 2 [−49 to −77], site 3 [−104 to −127], site 4 [−249 to −277], and site 5 [−214 to −247]) were used as probes to isolate putative transcription factors expressed as β-galactosidase fusion proteins. Only one of the clones which was isolated from screening with a site 4 probe will be discussed in this report; sites 1 and 2, which displayed little homology to site 4, served as nonspecific probes. A human fetal brain-derived λgt11 cDNA library (22 weeks of gestation; Clontech) was plated on nitrocellulose filters and screened with a labeled, multimerized DNA-binding site according to the method of Vinson et al. (39). Catenates of approximately 200 bp, produced by incubation of an oligonucleotide with T4 DNA ligase, were gel purified and labeled with [^{32}P]dCTP (Amersham) by nick translation

* Corresponding author.

(Bethesda Research Laboratories). The sequence of the oligonucleotide used was as follows:

AAGGATCAGTTGGAAGTTTCCAGGACATCTTC
TAGTCAACCTTCAAAGGTCCTGTAGAAGTTCC (site 4)

Purified phage plaques were screened with specific and nonspecific probes. Nonspecific probes (site 1 and site 2), which corresponded to regions of the PLP promoter other than the site 4 (2), had the following sequences:

TTAAGGGGGTTGGCTGTCAATCAGAAA
CCCCCAACCGACAGTTAGTCTTTAATT (site 1)

GGAAAAGGGGAGGAGAAGGGGAGGAG
TTCCCCTCCTCTTCCCCTCCTCCCTT (site 2)

The presence of β-galactosidase fusion proteins was determined on immunoblots of cell lysates from uninduced or isopropylthiogalactopyranoside (IPTG)-induced (0.1 mM IPTG for 3 h) cultures, using a peroxidase staining (Vectastain ABC kit; Vector Laboratories) of anti-β-galactosidase antibody (5 Prime→3 Prime, Inc.). After phage plaque purification, the cDNA inserts were subcloned into vector pT7T3 18U (Pharmacia) for sequence determination by the dideoxynucleotide chain termination method on double-stranded DNA templates with Sequenase (United States Biochemical).

For the isolation of overlapping cDNAs, a 444-bp *Eco*RI-*Eco*RV fragment corresponding to the 5′ end of the initial clone (MyTIa) was labeled by nick translation and used to screen a fetal brain λgt11 library (17 to 18 weeks of gestation; Stratagene) by plaque hybridization (10). pBluescript plasmids containing the cDNA inserts were excised from the λZapII vector according to the supplier (Stratagene).

Expression of recombinant protein in *Escherichia coli*. Three different constructs containing portions of the MyTI sequence were cloned into the bacteriophage T7 expression vector pET3a (37). From this vector, inserted coding sequences are translated as a fusion protein in which the first 12 amino acids are encoded by the phage gene *10* and linker sequences. A 622-amino-acid fusion protein, which included the downstream cluster of four zinc fingers and flanking regions, was prepared by ligating the *Bam*HI-*Bgl*II fragment (amino acids 116 to 726 in Fig. 1) from the recombinant phage MyTIa with pET3a. A construct containing only the four fingers of MyTIa with minimal flanking regions (amino acids 373 to 603; nucleotides 1117 to 1808 in Fig. 1) was cloned into pET3a by a polymerase chain reaction (PCR)-based strategy, as was a construct containing only the upstream two fingers and short flanking regions encoded by MyTIb (amino acids 1 to 122; nucleotides 1 to 366 in Fig. 1). The recombinant plasmids or the vector alone was introduced into *E. coli* BL21(DE3), a lysogenic strain harboring the T7 RNA polymerase gene under the control of the IPTG-inducible promoter *lac*UV5. Cells were grown in LB broth containing ampicillin (100 μg/ml) to an optical density at 600 nm of 0.5. IPTG was added to a final concentration of 0.4 mM, and incubation was continued at 37°C for 3 h. Bacterial extracts were prepared as described by Singh et al. (36).

Electrophoretic mobility shift analysis. The site 4 oligonucleotide end labeled with ^{32}P by T4 polynucleotide kinase (34) was incubated with extracts of *E. coli* expressing MyTI (10 μg of protein) and 4 μg of poly(dI-dC) (Pharmacia) in 20 μl of buffer containing 10 mM Tris-HCl (pH 8.0), 50 mM KCl, 0.5 mM EDTA, 1 mM dithiothreitol, and 5% glycerol. The mixture was incubated on ice for 10 min with or without a 100-fold molar excess of an unlabeled competitor and for 20 min with the labeled oligonucleotide. The DNA-protein complexes formed were resolved in a 6% polyacrylamide native gel containing 25 mM Tris-HCl (pH 8.3), 190 mM glycine, and 1 mM EDTA. The gel was then dried and autoradiographed.

For determining the metal requirements of MyTI, *E. coli* extracts were preincubated as indicated above with 0 to 5 mM zinc chloride, cobalt chloride, copper chloride, magnesium chloride, or manganese chloride (all 99.999% pure; Aldrich) or 1,10-phenanthroline (Sigma).

Southwestern (DNA-protein) analysis. The assay was carried out as described by Michael et al. (24). *E. coli* extracts were resolved by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis and electrophoretically transferred onto a nitrocellulose membrane. Immediately after transfer, the filter was washed with three changes of buffer A (10 mM Tris-HCl [pH 7.5], 5% nonfat dry milk, 10% glycerol, 2.5% Nonidet P-40, 0.1 mM dithiothreitol, 150 mM NaCl) for 3 h. The filter was briefly rinsed in buffer B (10 mM Tris-HCl [pH 7.5], 0.125% nonfat dry milk, 8% glycerol, 1 mM EDTA, 1 mM dithiothreitol, 40 mM NaCl) and hybridized with 5 ml of buffer B containing 5 mM $MgCl_2$, 50 μg of poly(dI-dC) per ml, and ^{32}P-labeled probe DNA. After incubation overnight at room temperature, the filter was washed three times with 10 mM Tris-HCl (pH 7.5)–50 mM NaCl and subjected to autoradiography. Proteins were visualized by staining with 10% Ponceau S (Sigma).

Northern (RNA) blot analysis. Total RNA was prepared from various tissues by the guanidinium thiocyanate-cesium chloride method essentially as described by Sambrook et al. (34). Poly(A)$^+$ RNA was isolated by using a poly(A) Quik mRNA purification kit (Stratagene). RNA was electrophoresed on a formaldehyde agarose gel, transferred to nylon membranes, and hybridized with the MyTIa insert uniformly labeled with [^{32}P]dCTP by a random-priming kit (Bethesda Research Laboratories). Blots were washed at room temperature in 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) and then washed twice with 0.1× SSC–1% SDS at 60°C for 30 min each time.

RT-PCR analysis of alternatively spliced transcripts. PCR was carried out with *Taq* polymerase as specified by the manufacturer (Perkin-Elmer/Cetus). For reverse transcriptase-linked PCR (RT-PCR), total RNA was annealed with oligo(dT)$_{12-18}$ and incubated with Moloney murine leukemia virus reverse transcriptase (Stratagene) to synthesize cDNA. The cDNA formed was subjected to 40 cycles of PCR amplification, using two sets of primers (oligonucleotide positions are shown in parentheses; refer to the sequence of Fig. 1 for comparison): 20-6 (5′-CGATGCT CAGGTTTTTGGCAAACGCATGCT-3′; bases 501 to 530) plus 20-10 (5′-TGGCTCTGACTCCTCAGGTTCCTCCTCT CT-3′; bases 1033 to 1062) and S2 (5′-CCCCTGGACACTG TGCGGAAGAGTT-3′; bases −14 to +10) plus S3 (5′-GGC CTGAGGATCCGATCGGAATTGG-3′; bases 341 to 365). Each cycle entailed 1 min at 94°C, 2 min at 60°C, and 3 min at 72°C. The PCR products were analyzed by agarose gel electrophoresis and ethidium bromide staining.

Of the two primers used in an RT-PCR assay to confirm the composite MyTI sequence, the 5′ primer originated from the MyTIa sequence (S1 [5′-CCAACACGTCGCTAGTCT TGAACTT-3′; bases 1783 to 1807]), and the 3′ primer corresponding to the MyTIb sequence was the S2 primer shown above.

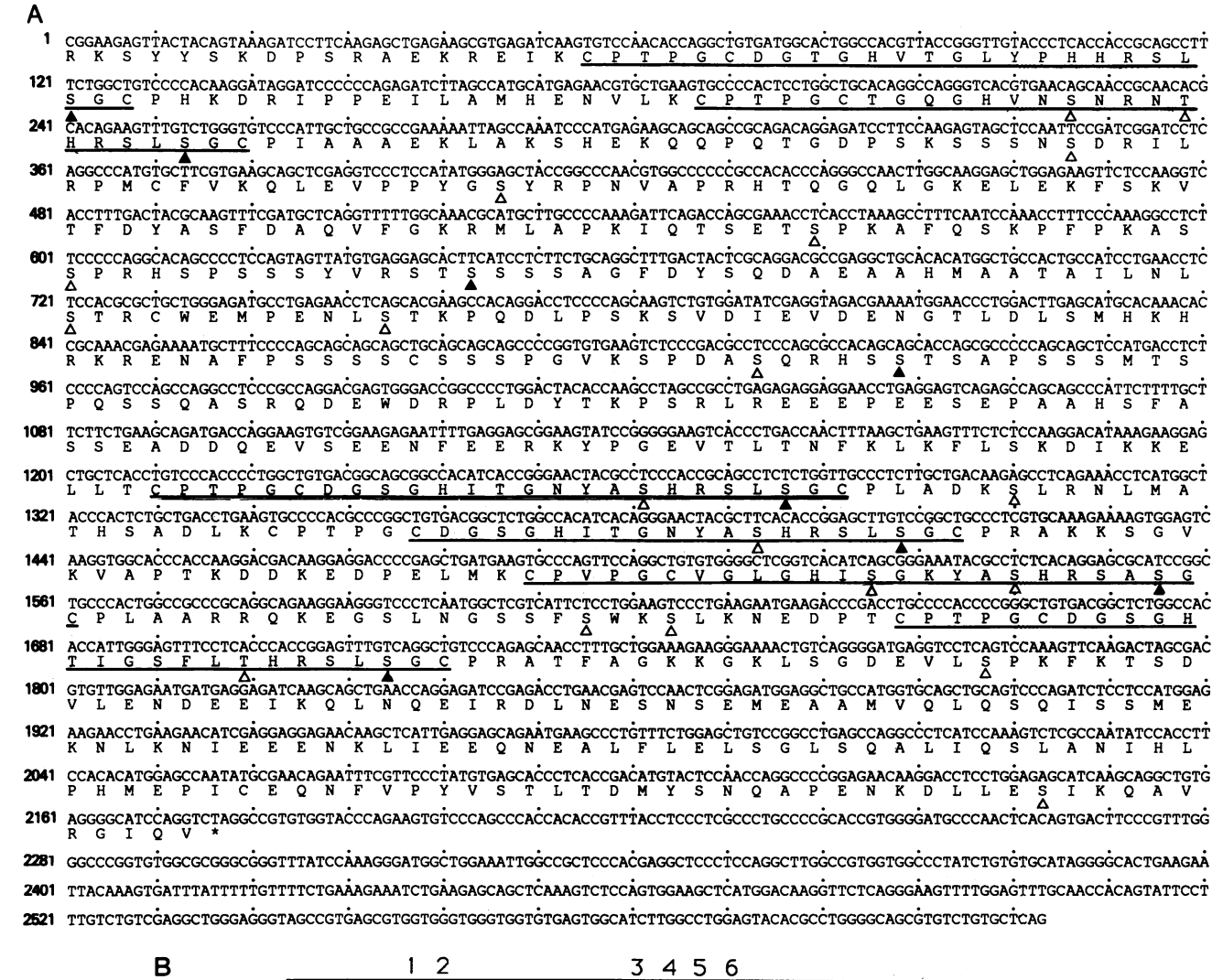

FIG. 1. (A) Nucleotide and predicted amino acid sequences from the MyTI cDNAs. The zinc finger regions of the amino acid sequence are underlined. The protein kinase C consensus sites (Ser/Thr-X-Lys/Arg [15]) are indicated by open triangles, and the Ca^{2+}-calmodulin-dependent kinase II sites (15) are represented by filled triangles. (B) Diagram of MyTI cDNA clones, in which the six zinc finger regions are numbered and shaded within the open reading frame (open box). The 3′ untranslated region is represented by the filled box. MyTIa extends from bases 347 to 2654, and MyTIb extends from bases 1 to 596.

Nucleotide sequence accession number. The GenBank nucleotide sequence accession number for the reported sequence is M96980.

RESULTS

Identification and sequence of a cDNA encoding a putative myelin transcription factor. To isolate cDNA clones encoding proteins that bind *cis* elements of the human PLP gene, a fetal brain expression library was screened with catenated oligonucleotides. From approximately 1 million phage plaques, several phage clones that exhibited binding to a recognition site (site 4) identified in previous work (2) were isolated. One positive clone, named MyTIa, had the largest insert size of 2.2 kb. On immunoblots, the β-galactosidase–MyTIa fusion protein was approximately 200,000 kDa, which indicated that MyTIa encoded a protein of 86 kDa. The DNA-binding specificity of this clone was examined by the filter binding assay. The MyTIa phage protein bound strongly to the site 4 probe but failed to bind the unrelated site 1 and site 2 probes.

MyTIa included the carboxy but not the amino terminus of the protein. To obtain larger and overlapping cDNAs, the cDNA segment from MyTIa was used to probe another fetal brain cDNA library. One recombinant phage (MyTIb) contained an overlapping cDNA insert that extended further but

FIG. 2. (A) Comparison of zinc finger domains of MyTI. For the consensus, amino acids common to at least four of the six fingers are shown at the bottom. (B) Schematic of a zinc finger, illustrating the four amino acids (C_2-HC) involved in metal coordination that participate in forming the knuckle of the finger.

still did not contain the 5′ end of the transcript. The MyTI nucleotide sequence determined from the overlapping cDNA clones and the deduced amino acid sequence are shown in Fig. 1A. To confirm that this composite sequence exists in vivo, primers corresponding to the ends of the sequence were used in an RT-PCR assay to amplify the expected 1,821-bp fragment.

Analysis of the predicted open reading frame encoding 726 amino acids reveals a total of six putative zinc fingers with a rather unusual structure in which three cysteines and a histidine would serve to coordinate metal binding (Fig. 2B). The fingers are present in two widely separated sets. One set (fingers 1 and 2) is located near the amino terminus, whereas the other set (fingers 3 to 6) is 316 amino acids away, near the carboxy terminus. While each finger is over 80% homologous to the other fingers, regions between fingers are not conserved. All of the fingers contain a constant spacing of 12 amino acids between the second cysteine and the histidine residue. The two stretches that form the "knuckle" of the finger (Fig. 2) are nearly identical among fingers. Thus, the six fingers in MyTI display extensive homology, suggesting that the repeats were created by successive duplication events.

Other features of the MyTI sequence include an acidic domain (45% glutamic or aspartic acid residues from bases 1037 to 1129 in Fig. 1) and serine-rich domains (58% from bases 598 to 655 and 48% from bases 866 to 982), both of which occur in the region between the two sets of zinc fingers. In addition, MyTI has 19 sites that fit the consensus for phosphorylation by protein kinase C and 8 potential Ca^{2+}-calmodulin-dependent kinase II sites (Fig. 2); interestingly, each of the zinc fingers except the first contains both a protein kinase C and a Ca^{2+}-calmodulin-dependent kinase II phosphorylation site.

A search of the sequences deposited in the GenBank, EMBL, and zinc finger (compiled by George Michaels) data bases did not reveal any entries with similarity to the MyTI protein. Southern blot analysis of human genomic DNA with MyTI cDNA as a probe suggests that there is a single MyTI gene in the human genome, as a simple pattern with few bands was detected in genomic DNA restricted with *Eco*RI, *Bam*HI, or *Not*I (data not shown). Hybridization under reduced stringency revealed many additional bands on a genomic Southern blot, suggesting that the MyTI gene may be a member of a large family of transcription factors that employ a Cys-X_5-Cys-X_{12}-His-X_4-Cys (C_2-HC) zinc finger motif to recognize DNA.

Specificity of MyTI binding. To avoid any possible interference of the β-galactosidase domain with DNA binding, MyTI cDNA containing two or four fingers was subcloned into another expression vector in which only the first 12 amino acids are foreign. To characterize the DNA-binding specificity of the protein segments encoded by MyTI, bacterial extracts were tested in gel mobility shift assays with the double-stranded site 4 oligonucleotide. In the presence of bacterial extract containing a four-finger segment of MyTI, site 4 formed three closely migrating, specific complexes (Fig. 3). That these complexes are related is suggested by the observation that freshly prepared extracts appear to contain only the most retarded band. Formation of specific complexes was abolished when extract was preincubated with a 100-fold excess of unlabeled site 4 oligonucleotide, whereas under identical conditions, addition of 100-fold molar excesses of oligonucleotides corresponding to two other sites in the PLP promoter (site 1 or site 2 oligonucleotide) had no noticeable effect on specific complex formation (Fig. 3, lanes 5 and 6).

The two-finger protein encoded by MyTIb displayed the same DNA-binding properties as did the four-finger protein encoded by MyTIa, as shown by a Southwestern analysis (Fig. 4). Binding of the site 4 oligonucleotide was observed for the upstream two zinc fingers (lane 1), the downstream four zinc fingers (lane 2), and the 610-amino-acid construct which encompassed the downstream four zinc fingers and flanking regions (lane 3). Thus, the minimal DNA-binding domain of MyTI is a two-finger unit, in this case the amino-terminal set of two fingers; the flanking regions, which diverge considerably between the two-finger and four-finger constructs, do not appreciably affect binding.

No binding proteins were detectable when as much as 50 μg of nuclear extract from mouse or human brain was analyzed by Southwestern blots (Fig. 4, lanes 4 to 6), perhaps because the proteins are present in relatively low amounts in these extracts.

Metal requirements of MyTI for DNA binding. Incubation of the MyTI protein with increasing concentrations of the zinc-chelating agent 1,10-orthophenanthroline impaired the formation of protein-DNA complexes, as assayed by gel mobility shifts (Fig. 5A). These results suggest that Zn^{2+} is required for DNA binding. Bacterially produced MyTI is probably maximally bound with zinc, as additional zinc did not increase the amount of specific complexes (compare 0 and 0.1 mM zinc chloride in Fig. 5A). Surprisingly, high concentrations of zinc were strongly inhibitory to complex

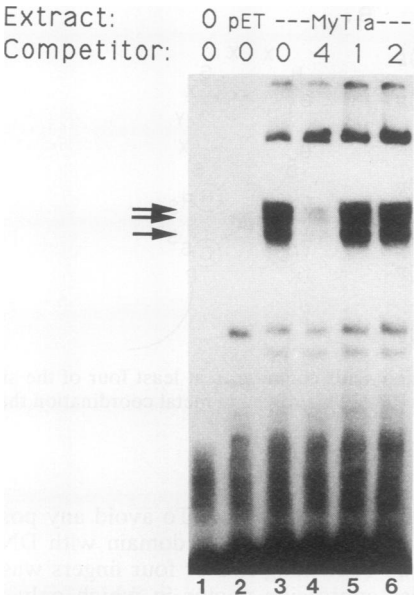

FIG. 3. DNA-binding specificity of MyTI. Bacterial extracts containing the MyTIa protein, a 610-amino-acid protein which includes the cluster of four zinc fingers, were tested in gel mobility shift assays using a ^{32}P-labeled site 4 oligonucleotide as a probe. The control extracts were prepared from bacteria carrying the vector without inserts (pET). Unlabeled competitor oligonucleotides (site 1, 2, or 4) were added at a 100-fold molar excess. Arrows mark the specific bands, which were present only in the MyTIa extract and nearly eliminated by an excess of unlabeled site 4 but were not detectably affected by unrelated oligonucleotides (site 1 or 2). The lower band represents nonspecific DNA binding, since the same complex was formed with the control extracts prepared from bacteria carrying the vector without insert.

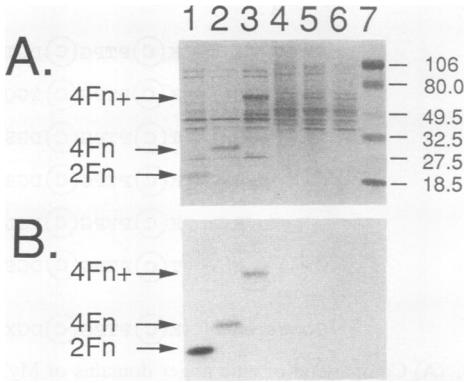

FIG. 4. Southwestern analysis of MyTI. Bacterial extracts containing one of three fusion proteins were electrophoresed, transblotted, and probed with a labeled DNA-binding site. The three fusion proteins tested for binding to the site 4 oligonucleotide included 2Fn, the upstream two zinc fingers in a 122-amino-acid segment encoded by MyTIb (lane 1), 4Fn, the downstream four fingers subcloned as a 230-amino-acid segment from MyTIa (lane 2), and 4Fn$^+$, the entire MyTIa coding region of 610 amino acids that encompasses the four-finger region (lane 3). See Materials and Methods for a complete description of the three zinc finger constructs. Nuclear extracts prepared from brain included 3-week-old mouse (lane 4), 22-year-old human (lane 5), and 8-month-old human (lane 6), prepared as previously described (2). Approximately 30 to 50 μg of protein was loaded per lane. Lane 7, low-range molecular weight protein standards (Bio-Rad). Molecular sizes (in kilodaltons) are noted at the right. (A) Ponceau S stain of the protein blot; (B) autoradiogram of radiolabeled site 4 oligonucleotide bound to the different zinc finger constructs.

formation. This effect has been observed with other zinc finger proteins (41) and may reflect a tendency of zinc at high concentrations to coordinate additional cysteine and histidine residues in a manner that affects the conformation and consequently the DNA binding of the protein.

Several other divalent cations, such as Co^{2+}, Mn^{2+}, and Mg^{2+}, had no effect on DNA binding over a range of concentrations (0.01, 0.1, 0.5, 2, and 5 mM) in three independent experiments. At concentrations of 0.1 mM, Cu^{2+} and Cd^{2+} had no effect on DNA binding, but higher concentrations reduced binding (Fig. 5B). This result is similar to that observed with another zinc finger protein, TFIIIA (25), in which copper and cadmium appear to displace bound zinc, thereby interfering with complex formation.

Tissue-specific expression of the MyTI gene. On Northern blots, a major transcript of 5.5 kb is expressed at very low levels in the brain, with expression highest earlier in development, i.e., embryonic day 15 in the mouse and 8 months after birth in humans (Fig. 6, lanes 1 and 3). A minor transcript of 1.0 kb is present in human cells at both time points in equivalent amounts, and murine brain contains trace amounts of at least four other transcripts, a doublet around 4 kb and single bands of 2.0 and 1.7 kb. MyTI transcripts are undetectable in other tissues by Northern blots at both early and late stages. Spleen, kidney, lung, liver, and pancreas displayed no MyTI transcripts at the peak of myelination in mice (Fig. 6, lanes 5 to 9). Likewise, at embryonic day 19, these tissues, along with heart, intestine, muscle, and testis, lacked detectable MyTI transcripts (data not shown).

Gel shift analysis of nuclear extracts also indicates that the MyTI binding activity is present predominantly in the nervous system, as it was not detected in liver, kidney, spleen muscle, heart tissue, or HeLa cells but was present in brain tissue and glial cells lines (data not shown). The only tissue outside of the nervous system that displayed binding activity by gel mobility shift assays was the lung, and this activity may be unrelated to the MyTI protein.

Since the MyTI message is rare, we analyzed RNAs from different tissues by using RT-PCR. Two combinations of primers were used, one corresponding to the region encompassing the first two zinc fingers and the other in the middle of the protein. The two-finger region was expressed in all tissues examined (Fig. 7B). Likewise, the four-finger cluster was ubiquitously expressed (data not shown). However, the hinge region between the two zinc finger clusters showed a restricted tissue distribution, as no PCR products were detected in spleen, liver, and kidney (Fig. 7A). In these tissues, the proteins encoded by the MyTI gene would contain only one of the two clusters of zinc fingers. Thus, the expression of MyTI transcripts is greatly enriched but not confined to the nervous system, unless the PCR products of nonneural tissue represent leaky transcription. These results also indicate that in nonneural tissue, the products of the MyTI locus may be alternatively spliced.

DISCUSSION

To define the regulatory controls exerted on cells of the glial lineage, a selective binding procedure (17, 36) was used to clone a protein that specifically recognizes a myelin gene.

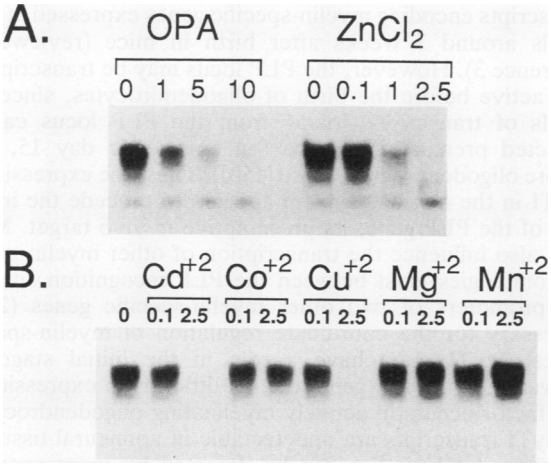

FIG. 5. Effects of metals on MyTI-DNA interaction. Bacterial extracts containing the MyTI protein were incubated in a standard DNA-binding reaction with the indicated concentrations (millimolar) of agents. (A) Increasing concentrations of the zinc chelator 1,10-orthophenanthroline (OPA) or of zinc chloride (Zn^{2+}) were tested. (B) Two concentrations (0.1 and 2.5 mM) of the divalent metals cadmium chloride (Cd^{2+}), cobalt chloride (Co^{2+}), copper chloride (Cu^{2+}), magnesium chloride (Mg^{2+}), and manganese chloride (Mn^{2+}) were used.

MyTI binds to a previously defined *cis* element of the human PLP promoter (2) by gel mobility shift and filter binding assays. MyTI contains six repeats of an unusual zinc finger domain, in which binding of the metal is coordinated by three cysteines and one histidine residue. The MyTI fingers are grouped into two widely separated clusters consisting of an upstream set of two fingers followed by a 319-amino-acid hinge region and then four adjacent fingers. As anticipated by the high degree of conservation between fingers, the upstream two fingers display DNA-binding properties simi-

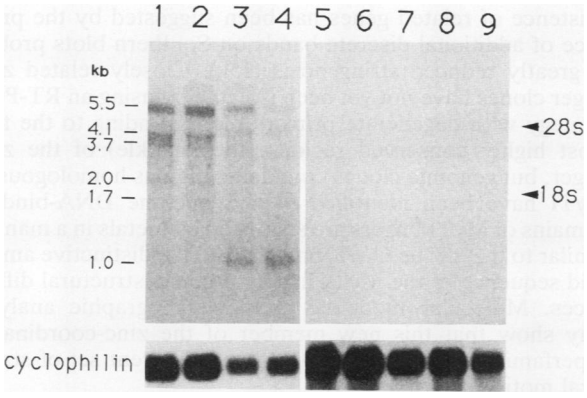

FIG. 6. Northern blot analysis of MyTI. The blot was hybridized with the 2.2-kb cDNA insert from MyTIa. The size of the major MyTI transcript is 5.5 kb. Positions of the 28S (4,718 bases) and 18S (1,874 bases) bands are indicated. Ten micrograms of poly(A)$^+$ RNA from each tissue was analyzed except for the lung and liver samples, which contained 5 μg of poly(A)$^+$ RNA. Lanes: 1, mouse brain (embryonic day 15); 2, mouse brain (5 days after birth); 3, human brain (8 months after birth); 4, human brain (22 years); 5 to 9, spleen, kidney, lung, liver, and pancreas, respectively, from 3-week old mice. Shown at the bottom is the same blot reprobed with a cyclophilin cDNA insert.

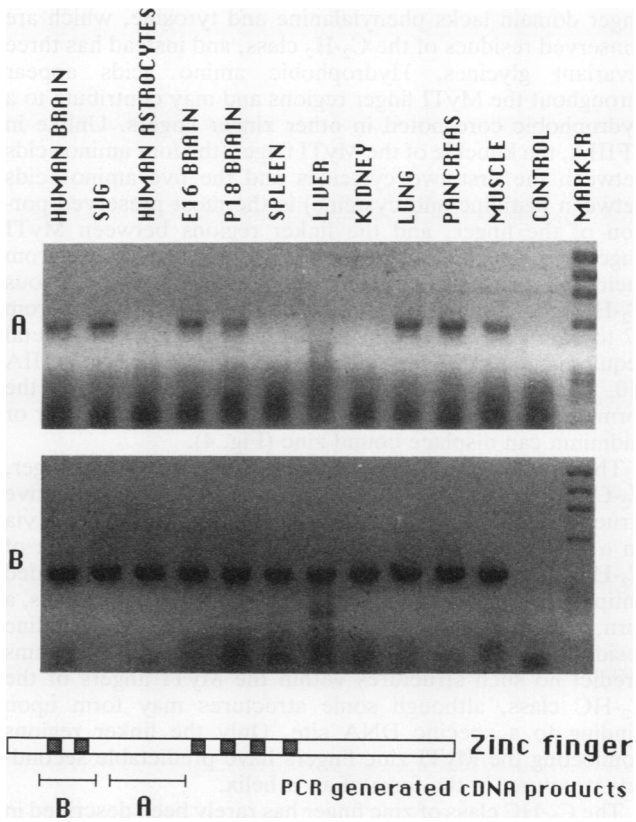

FIG. 7. Agarose gel electrophoresis of PCR products. ϕX174 DNA digested with *Hae*III was used as a size marker. The combinations of primers are 20-6-20-10 (A in the diagram) S2 and S3 (B in the diagram). Expected sizes of PCR-generated bands are 563 bp for the linker region (A) and 369 bp for the zinc finger region (B). The human astrocytes, a gift from Kei Amemiya and Eugene Major, were prepared from fetal brain of 16 weeks gestation and maintained in culture for 2 weeks. SVG is a cell line derived by Eugene Major from human fetal glial cells (21) and was cultured as previously described (2). Human brain is a sample taken 8 months after birth; E16 brain refers to embryonic day 16 mouse brain; P18 brain refers to postnatal day 18 mouse brain. The remaining tissues are from a 3-week-old mouse.

lar to those of the downstream four fingers. Two fingers of MyTI confer specificity, an emerging feature of other zinc finger proteins in which a single pair of fingers constitutes a minimal DNA recognition domain (14).

The use of zinc to stabilize peptide domains that interact with nucleic acid is a strategy adopted by many eukaryotic transcription factors. Three major groups of zinc-mediated peptide loops have been defined first by amino acid sequence comparisons and subsequently by crystallographic analysis (reviewed in references 1, 7, and 16). Members of the zinc finger superfamily are classified by the amino acids employed to coordinate metal ligands (usually zinc): either two cysteine and two histidine residues in a Cys-X_{2-5}-Cys-$X_{12,13}$-His-X_{2-5}-His arrangement (referred to as C_2-H_2) or all cysteines in a C_2-C_2 finger or in a C_x cluster. The C_2-HC fingers of MyTI are most similar to the C_2-H_2 class typified by the RNA polymerase III transcription factor TFIIIA (25), which has nine tandemly repeated sequences of approximately 30 amino acid residues. Both proteins have a 12-amino-acid loop between cysteine and histidine, although the amino acids that comprise the loop differ markedly. The MyTI zinc

finger domain lacks phenylalanine and tyrosine, which are conserved residues of the C_2-H_2 class, and instead has three invariant glycines. Hydrophobic amino acids appear throughout the MyTI finger regions and may contribute to a hydrophobic core noted in other zinger fingers. Unlike in TFIIIA, the knuckle of the MyTI finger (the four amino acids between the first two cysteines and the five amino acids between histidine and cysteine) is the most preserved portion of the finger, and the linker regions between MyTI fingers are the least conserved. These linkers differ from their well-conserved counterparts of the closely contiguous C_2-H_2 fingers in both their variable size, which ranges from 17 to 26 amino acids, and divergent sequences. The metal requirements for DNA binding are very similar for TFIIIA (10, 25) and MyTI. Zinc is a cofactor needed for the formation of DNA-protein complexes, and either copper or cadmium can displace bound zinc (Fig. 4).

The three classes of zinc-binding domains (C_2-H_2 finger, C_2-C_2 finger, and C_x cluster) have remarkably distinctive structures, although all contact the major groove of DNA via an α helix (20, 22, 30). The three-dimensional structure of C_2-H_2-type fingers (18, 29, 30) consists of a two-stranded antiparallel β sheet that includes the two cysteine residues, a turn, and then an α helix that includes the two histidine residues. The Chou-Fasman (4) and Garnier (9) algorithms predict no such structures within the MyTI fingers of the C_2-HC class, although some structures may form upon binding to a specific DNA site. Only the linker regions connecting the MyTI zinc fingers have predictable secondary structure, in the form of an α helix.

The C_2-HC class of zinc finger has rarely been described in other putative transcription factors and, when present, is found as an isolated finger in a protein containing many other conventional fingers, such as in Xfin (33), Evi-(26), and the *su(Hw)* protein (28). Whether these isolated C_2-HC fingers actually bind DNA is unknown but unlikely in view of functional studies of the C_2H_2 family suggesting that a pair of zinc fingers forms the fundamental sequence recognition unit (14). The C_2-HC motif has also been found in nucleic acid-binding proteins that display less sequence specificity than do transcription factors, including a poly(ADP-ribose) polymerase (13) and *gag*-derived retroviral proteins (5). However, in these cases the structure and composition of the zinc finger domain are much different than in MyTI, with a very large loop region for the poly(ADP-ribose) polymerase (28 to 30 amino acids) and a small one (4 amino acids) for the RNA-binding *gag* proteins. The structure of the human immunodeficiency virus-encoded *gag* protein determined from a nuclear magnetic resonance-based distance geometry approach suggests that this class of zinc-coordinated structures is folded very differently from those of the C_2-H_2, C_2-C_2, or C_x class (38). As proposed by Luisi, zinc may have been rediscovered several times in eukaryotes as a way of stabilizing nucleic acid-binding domains, and each reappearance of the zinc coordination scheme seems to have engendered a distinct folding pattern of the zinc module (19).

MyTI transcripts are expressed predominantly in the brain, with higher levels noted in early development. The cells in the brain that synthesize MyTI appear to be the progenitors of oligodendrocytes, as indicated by preliminary Northern analysis of cultures highly enriched in either progenitors, astrocytes, microglia, or oligodendrocytes (data not shown). These observations are consistent with the reduced levels of MyTI transcripts in adult brain, which also has much fewer progenitor cells. Myelination in both humans and rodents is primarily a postnatal process, with transcripts encoding myelin-specific genes expressed at peak levels around 3 weeks after birth in mice (reviewed in reference 3). However, the PLP locus may be transcriptionally active before the birth of oligodendrocytes, since low levels of transcripts arising from the PLP locus can be detected prenatally in mice (on embryonic day 15, long before oligodendrocytes arise [15b]). Thus, the expression of MyTI in the nervous system appears to precede the induction of the PLP gene, its presumptive in vivo target. MyTI may also influence the transcription of other myelin genes, as homologies exist between the PLP recognition site and the promoters of two other myelin-specific genes (2). If necessary for the coordinate regulation of myelin-specific genes, MyTI must have a role in the initial stages of activation of myelin gene loci, as little or no expression of this factor occurs in actively myelinating oligodendrocytes.

MyTI transcripts are undetectable in nonneural tissue by Northern analysis but can be detected by more sensitive methods (Fig. 7). Alternative splicing of MyTI transcripts was suggested by the presence of PCR products in liver, kidney, and spleen that correspond to the amino-terminal set of two tandem fingers and the absence in these tissues of PCR products that correspond to the adjacent hinge region of the protein. Precedent exists for alternative splicing of genes encoding zinc fingers, an event that can generate proteins with distinct binding specificities (31). If actually translated in nonneural tissue, the MyTI isoforms containing the amino-terminal two fingers, the carboxy-terminal four fingers, or all six fingers may differentially affect target genes. While both sets of fingers recognize the same DNA sequence by Southwestern analysis (Fig. 4), the in vivo binding affinities for target sites may differ. Another feature that may influence target specificity is the large spacing between the two sets of zinc fingers in the six-finger protein, which may facilitate the binding of distant or even unlinked recognition sites. By simultaneously binding both proximal and distal recognition sites, a zinc finger protein could be involved in the assembly and domain organization of higher-order DNA-protein complexes (8, 32).

MyTI may represent a zinc finger subfamily of transcription factors that coordinate metal by a C_2-HC structure. The existence of related genes has been suggested by the presence of additional discrete bands on Southern blots probed at greatly reduced stringencies (15a). Closely related zinc finger clones have not yet been isolated by using an RT-PCR strategy with degenerate primers corresponding to the two most highly conserved regions (the knuckle) of the zinc finger, but genomic clones containing regions homologous to MyTI have been identified (40). While the DNA-binding domains of MyTI fingers are stabilized by metals in a manner similar to that of the C_2-H_2 finger class, the distinctive amino acid sequence of the MyTI fingers predicts structural differences. Molecular modeling and crystallographic analysis may show that this new member of the zinc-coordinated superfamily interacts with DNA by a different set of structural motifs.

ACKNOWLEDGMENTS

We thank George Michaels for comparing MyTI against the zinc finger data base, F. W. Studier for the gift of the pET3a expression system, Kei Amemiya and Eugene Major for the gift of human fetal astrocytes, and Jean Wrathall and Eric Verdin for helpful comments on the manuscript.

J.G.K. was a fellow of the National Research Council and the National Multiple Sclerosis Foundation.

REFERENCES

1. **Berg, J. M.** 1990. Zinc fingers and other metal-binding domains. J. Biol. Chem. **265**:6513–6516.
2. **Berndt, J., J. G. Kim, and L. D. Hudson.** 1992. Identification of *cis*-regulatory elements in the myelin proteolipid protein (PLP) gene. J. Biol. Chem. **267**:14730–14737.
3. **Campagnoni, A. T., and W. B. Macklin.** 1988. Cellular and molecular aspects of myelin protein gene expression. Mol. Neurobiol. **2**:41–89.
4. **Chou, P. Y., and G. D. Fasman.** 1978. Empirical predictions of protein conformation. Annu. Rev. Biochem. **47**:251–276.
5. **Covey, S. N.** 1986. Amino acid sequence homology in gag region of reverse transcribing elements and the coat protein gene of cauliflower mosaic virus. Nucleic Acids Res. **14**:623–633.
6. **Dynan, W. S., and R. Tjian.** 1985. Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. Nature (London) **316**:774–778.
7. **Evans, R. M., and S. M. Hollenberg.** 1988. Zinc fingers: gilt by association. Cell **52**:1–3.
8. **Fan, C.-M., and T. Maniatis.** 1990. A DNA-binding protein containing two widely separated zinc finger motifs that recognize the same DNA sequence. Genes Dev. **4**:29–42.
9. **Garnier, J., D. J. Osguthorpe, and B. Robson.** 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. **120**:97–120.
10. **Hanas, J. S., D. J. Hazuda, D. F. Bogenhagen, F. Y. Wu, and C. W. Wu.** 1983. Xenopus transcription factor A requires zinc for binding to the 5S RNA gene. J. Biol. Chem. **258**:14120–14125.
11. **Hudson, L. D.** 1990. Molecular biology of myelin proteins. Semin. Neurosci. **2**:483–496.
12. **Huynh, T. V., R. A. Young, and R. W. Davis.** 1985. Construction and screening cDNA libraries in λgt10 and λgt11, p. 49–78. *In* D. M. Glover (ed.), DNA cloning—a practical approach, vol. 1. IRL Press, Oxford.
13. **Ikejima, M., S. Noguchi, R. Yamashita, T. Ogura, T. Sugimura, D. M. Gill, and M. Miwa.** 1990. The zinc fingers of human poly(ADP-ribose) polymerase are differentially required for the recognition of DNA breaks and nicks and the consequent enzyme activation. J. Biol. Chem. **265**:21907–21913.
14. **Keller, A. D., and T. Maniatis.** 1992. Only two of the five zinc fingers of the eukaryotic transcriptional repressor PRDI-BF1 are required for sequence-specific DNA binding. Mol. Cell. Biol. **12**:1940–1949.
15. **Kennelly, P. J., and E. G. Krebs.** 1991. Consensus sequences as substrate specificity determinants for protein kinases and protein phosphatases. J. Biol. Chem. **266**:15555–15558.

15a. **Kim, J.** Unpublished data.

15b. **Kim, J., and L. Hudson.** Unpublished data.

16. **Klug, A., and D. Rhodes.** 1987. "Zinc fingers': a novel protein motif for nucleic acid recognition. Trends. Biochem. Sci. **12**: 464–469.
17. **Landschulz, W. H., P. F. Johnson, E. Y. Adashi, B. J. Graves, and S. L. McKnight.** 1988. Isolation of a recombinant copy of the gene encoding C/EBP. Genes Dev. **2**:786–800.
18. **Lee, M. S., G. P. Gippert, K. W. Soman, D. A. Case, and P. E. Wright.** 1989. Three-dimensional structure of a single zinc finger DNA-binding domain. Science **245**:635–637.
19. **Luisi, B. F.** 1992. Zinc standard for economy. Nature (London) **356**:379–380.
20. **Luisi, B. F., W. X. Xu, Z. Otwinowski, L. P. Freedman, K. R. Yamamoto, and P. B. Sigler.** 1991. Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. Nature (London) **352**:497–505.
21. **Major, E., A. Miller, P. Mourrain, R. Traub, E. De Widt, and J. Sever.** 1985. Establishment of a line of human fetal glial cells that supports a JC virus multiplication. Proc. Natl. Acad. Sci. USA **82**:1257–1261.
22. **Marmorstein, R., M. Carey, M. Ptashne, and S. C. Harrison.** 1992. DNA recognition by GAL4: structure of a protein-DNA complex. Nature (London) **356**:408–414.
23. **McKnight, S., and R. Tjian.** 1986. Transcriptional selectivity of viral genes in mammalian cells. Cell **46**:795–805.
24. **Michael, N., D. Spector, P. Mavromara-Nazos, T. M. Kristie, and B. Roizman.** 1988. The DNA-binding properties of the major regulatory protein α4 of herpes simplex viruses. Science **239**: 1531–1534.
25. **Miller, J., A. D. McLachlan, and A. Klug.** 1985. Repetitive zinc-binding domains in the protein transcription factor TFIIIA from Xenopus oocytes. EMBO J. **4**:1609–1614.
26. **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.
27. **Nave, K.-A., and G. Lemke.** 1991. Induction of the myelin proteolipid protein (PLP) gene in C6 glioblastoma cells: functional analysis of the PLP promoter. J. Neurosci. **11**:3060–3069.
28. **Parkhurst, S. M., D. A. Harrison, M. P. Remington, C. Spana, R. L. Kelley, R. S. Coyne, and V. G. Corces.** 1988. The Drosophila *su(Hw)* gene, which controls the phenotypic effect of the gypsy transposable element, encodes a putative DNA-binding protein. Genes Dev. **2**:1205–1215.
29. **Parraga, G., S. J. Horvath, A. Eisen, W. E. Taylor, L. Hood, E. T. Young, and R. E. Klevit.** 1989. Zinc-dependent structure of a single-finger domain of yeast ADR1. Science **241**:1489–1492.
30. **Pavletich, N. P., and C. O. Pabo.** 1991. Zinc finger-DNA recognition: crystal structure of a zif 268-DNA complex at 2.1 Å. Science **252**:809–817.
31. **Read, D., and J. L. Manley.** 1992. Alternatively spliced transcripts of the Drosophila tramtrack gene encode zinc finger proteins with distinct DNA binding specificities. EMBO J. **11**:1035–1044.
32. **Reuter, G., M. Giarre, J. Farah, J. Gausz, A. Spierer, and P. Spierer.** 1990. Dependence of position-effect variegation in Drosophila on dose of a gene encoding an unusual zinc-finger protein. Nature (London) **344**:219–223.
33. **Ruiz i Altaba, A., H. Perry-O'Keefe, and D. A. Melton.** 1987. Xfin: an embryonic gene encoding a multifingered protein in Xenopus. EMBO J. **6**:3065–3070.
34. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
35. **Shea, M. J., D. L. King, M. J. Conboy, B. D. Mariani, and F. C. Kafatos.** 1990. Proteins that bind to Drosophila chorion cis-regulatory elements: a new C_2H_2 zinc finger protein and a C_2C_2 steroid receptor-like component. Genes Dev. **4**:1128–1140.
36. **Singh, H., J. H. LeBowitz, A. S. Baldwin, Jr., and P. A. Sharp.** 1988. Molecular cloning of an enhancer binding protein: isolation by screening of an expression library with a recognition site DNA. Cell **52**:415–423.
37. **Studier, F. W., and B. Moffatt.** 1986. Use of bacteriophage T7 polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189**:113–130.
38. **Summers, M. F., T. L. South, B. Kim, and D. R. Hare.** 1990. High-resolution structure of an HIV zinc fingerlike domain via a new NMR-based distance geometry approach. Biochemistry **29**:329–340.
39. **Vinson, C. R., K. L. LaMarco, P. F. Johnson, W. H. Landschulz, and S. L. McKnight.** 1988. In situ detection of sequence-specific DNA binding activity specified by a recombinant bacteriophage. Genes Dev. **2**:801–806.
40. **Wrathall, J., C. Wiese, and L. Hudson.** Unpublished data.
41. **Zabel, U., R. Schreck, and P. A. Baeuerle.** 1991. DNA binding of purified transcription factor NF-κB. Affinity, specificity, Zn^{2+} dependence, and differential half-site recognition. J. Biol. Chem. **266**:252–260.