**Nuclear Protein Binding to the 5′ Enhancer Region of the Intracisternal A Particle Long Terminal Repeat***

(Received for publication, February 28, 1992)

Michael Zierler†, Robert J. Christy‡, and Ru Chih C. Huang§

From the Department of Biology, The Johns Hopkins University, Baltimore, Maryland 21218

The interactions of nuclear proteins from embryonal carcinoma cells (PCC3) with the long terminal repeats (LTRs) of murine intracisternal A particle (IAP) genes were studied. Two protein-DNA complexes were detected between PCC3 nuclear extract and IAP LTRs in a gel mobility shift assay. An additional complex was observed when enriched fractions from a heparin-agarose column were used as the source of proteins. Two regions within the LTR of IAP 81 were identified as the sites of protein interaction by DNase I protection. One region encompassed 43 nucleotides within the U3 region at the 5′ end of LTR 81. The other covers a 78 base pair region lying within 100 nucleotides upstream from the transcription initiation site. Studies using constructs containing intact or deleted versions of the LTR fused to the bacterial chloramphenicol acetyltransferase gene indicated that the absence of the 5′ 47 base pairs reduced the level of chloramphenicol acetyltransferase transcription to 20% of that driven by the entire LTR. Southwestern analysis of PCC3 nuclear extracts and column fractions revealed that a 28,000- and a 46,000-dalton protein were the major species that interact with the 5′ end of IAP LTR 81.

Intracisternal A particle (IAP) genes comprise a large family of retrovirus-like elements (or retrotransposons) that are defective in envelope protein synthesis and exist without a viral envelope. The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) X01705.†

‡ Supported by the National Research Service Predoctoral Training Grant 5 T33 GM07231 from the National Institutes of Health. Present address: Dept. of Microbiology, State University of New York, Stony Brook, New York 11794.

§ Supported by the National Institutes of Health. Present address: Institute of Biotechnology Center for Molecular Medicine, The University of Texas Health Science Center at San Antonio, San Antonio, Texas 78245.

† To whom correspondence should be addressed: Tel.: 410-274-1518; Fax: 410-617-7735.

Supported by United States Public Health Service Research Grant 5 RO1 CA48263. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked in accordance with 18 U.S.C. Section 1734 solely to indicate this fact. The abbreviations used are: IAP, intracisternal A particle; LTR, long terminal repeat(s); MIPF gene, mouse IAP promoted placental gene; CAT, chloramphenicol acetyltransferase; DMEM, Dulbecco's modified Eagle's medium; HEPES, 4-(2-hydroxyethyl)-1-piperazinethanesulfonic acid; DTT, dithiothreitol; PMSF, phenylmethylsulfonyl fluoride; SDS, sodium dodecyl sulfate; bp, base pair; ATF, activation transcription factor; CRE, cAMP-response element.

approximately 1000 IAP retrotransposons/haploid genome (1–3). IAP particle formation and the expression of IAP genes are most often seen in cells that are actively dividing, such as immature thymocytes, tumor cells, and embryos (4–6). They have also been seen, rarely but consistently, in a variety of normal mouse tissues, including spleen, muscle, pancreas, ovary, thymus, and placenta as well as macrophage and fibroblast cells (7–10). IAP particles and mRNA are frequently seen in mouse precancerous and transformed cells and are especially abundant in established tumor cell lines. These include teratocarcinoma-derived cell lines, the embryonal carcinoma (EC) cell lines, PCC3, PCC4, and PCC6, and the parietal yolk sac cell line, PYS-2 (11, 12). As with retroviral proviruses, the IAP retrotransposons are composed of several protein-coding sequences flanked by long terminal repeats (LTR). The transcription initiation site is located at the boundary of the U3 and R regions of the LTR that is 5′ to the coding sequence (13–15). This LTR also contains all of the cis-acting elements required for the promotion and enhancement of IAP gene transcription (15, 16). Throughout the more than 20 sequenced LTRs, isolated from IAP clones and transposed elements, are a variety of sequences similar to recognition sites for eukaryotic transcriptional activators. These include sites for Sp1, AP1, and CCAAT enhancer binding protein (23, 43), a glucocorticoid-responsive element (38, 39), an enhancer core homology (13, 40), a CAMP-responsive element (41), a CCAAT box (42), and a TATA box (44).

In our laboratory, we have analyzed the ability of several LTRs, from IAP genomic clones and cDNA clones, to control the expression of the bacterial chloramphenicol acetyltransferase (CAT) gene transfected into PCC3-A/1 cells (15, 17). One LTR, from the genomic clone IAP 81C, has been dissected to determine the roles that different regions within the LTR play in regulating gene transcription (15). Due to the success of this work, and because it has been shown that undifferentiated PCC3 cells contain high levels of endogenous IAP mRNA (18, 19), we grew large quantities (>10⁵) of undifferentiated PCC3 cells, and isolated and characterized several nuclear proteins that interact with transcription regulatory elements of IAP LTR 81.

MATERIALS AND METHODS

Cell Culture

The murine embryonal carcinoma cell line, PCC3-A/1, was maintained in DMEM, supplemented with 10% fetal bovine serum, 4 g/liter glucose, 50 units/ml penicillin, and 50 mg/ml streptomycin. Cells were fed every 24–36 h and passaged every 48–72 h.

For large scale cultures, (10⁶ cells), cells were started in 180-cm² dishes and then transferred to roller bottles. PCC3 cells have a tendency to stick together, forming clumps on the tissue culture plastic or failing to adhere to the surface. To maximize PCC3 cell viability and growth in roller bottles, a strict protocol was followed. Each roller bottle was pretreated overnight with 1% gelatin in calcium and magnesium-free phosphate-buffered saline. 15 ml of gelatin/FBS was added to each 1000 ml roller bottle.
was added to each 850-cm² roller bottle, and rolled at 0.2 rpm for 18 h. After treatment, the bottles were used immediately or stored at 4 °C. Prior to adding the PCC3 cells, the excess gelatin was removed, 150-200 ml of supplemented DMEM was added, and the bottles were placed, with loosened caps, into the incubator for several hours, allowing the medium to be warmed and equilibrate. PCC3 cells were added to the roller bottles at 5 × 10⁵ cells/850-cm² bottle. The bottles were placed in the CO₂ incubator for 30 min, rotating the bottles a quarter of a turn every 5-10 min. (This last step helped ensure that the cells were dispersed evenly over the surface of the bottle.) Finally, the caps were tightened, and the bottles were rolled at 0.2 rpm at 37 °C. The cells were fed every 2 days by replacing 100 ml of medium with fresh supplemented DMEM. A bottle seeded with 5 × 10⁵ cells yielded 5 × 10⁶ cells in 4-5 days. For Southwestern assays, small scale cultures of several other cell lines were grown. Human HeLa cells, mouse myeloma cell line MOPC-315, mouse fibroblasts LTK-, and undifferentiated PCC3 cells were grown in 180-cm² dishes as described (19). The cells were induced by centrifugation.

**Preparation of Nuclear Extracts**

All nuclear extracts were prepared as described by Dignam et al. (20) with the following modifications. Cells were harvested in RPMI by scraping with a rubber policeman. The cells were centrifuged at 1000 × g for 12 min and washed once with cold calcium and magnesium-free phosphate-buffered saline. The cell pellet was resuspended in five times its volume in 10 mM HEPES (pH 7.6 at 4 °C), 1.5 mM MgCl₂, and 10 mM KCl, incubated on ice for 10 min, and centrifuged at 1000 × g for 10 min. The pellet was resuspended in twice the original packed cell volume with 20 mM HEPES (pH 7.6), 250 mM sucrose, 25 mM KCl, 2 mM MgCl₂, 3 mM CaCl₂, and 0.1% Triton X-100. The cells were lysed with 5-10 strokes of an pestle in an all glass homogenizer. (Lysis was checked by phase microscopy.) The nuclei were collected at 1500 × g for 20 min and washed once in the same volume of buffer minus the Triton X-100 and collected again by centrifugation.

The crude nuclei were resuspended in 3 ml/10⁸ cells with 20 mM HEPES (pH 7.6), 25% glycerol, 2 mM MgCl₂, 0.2 mM EDTA, 0.5 mM DTT, 20% glycerol, 1 mM PMSF, and transferred to Corex tubes. The nuclei were centrifuged at 1000 × g for 12 min and washed once with cold calcium and magnesium-free phosphate-buffered saline. The cell pellet was resuspended in five times its volume in 10 mM HEPES (pH 7.6 at 4 °C), 1.5 mM MgCl₂, and 10 mM KCl, incubated on ice for 10 min, and centrifuged at 1000 × g for 10 min. The pellet was resuspended in twice the original packed cell volume with 20 mM HEPES (pH 7.6), 250 mM sucrose, 25 mM KCl, 2 mM MgCl₂, 3 mM CaCl₂, and 0.1% Triton X-100. The cells were lysed with 5-10 strokes of an pestle in an all glass homogenizer. (Lysis was checked by phase microscopy.) The nuclei were collected at 1500 × g for 10 min, aliquoted, and stored in liquid N₂. Protein concentrations were 2-4 mg/ml as previously described (19).

**Column Chromatography**

DEAE-Cellulose—PCC3 nuclear extract (45 ml, 72 mg), adjusted to 100 mM KCl, was loaded onto a DEAE-cellulose column (50 ml, 2.4 × 11 cm) that had been equilibrated with 100 mM KCl in DEAE buffer (50 ml HEPES (pH 7.5), 2 mM DTT, 0.5 mM PMSF, 20% glycerol, 1 mM MgCl₂). The column was washed with 80 ml of 100 mM KCl in DEAE buffer, followed by a stepped increase to 200 mM KCl in DEAE buffer (90 ml), and finally eluted with 1 M KCl in DEAE buffer (50 ml).

**Heparin-Agarose—Fractions 4–9 from the DEAE-cellulose column were pooled, concentrated, and adjusted to 100 mM KCl. The protein (7.5 ml, 16 mg) was loaded onto a heparin-agarose column (8 ml, 0.74 × 2.5 cm) that had been equilibrated with 100 mM KCl in TM buffer (50 mM Tris-HCl (pH 7.9), 10 mM MgCl₂, 5 mM DTT, 3.5 mM PMSF, 1 mM EDTA, 15% glycerol). The column was washed with 25 ml of 100 mM KCl in TM buffer, followed by elution with 25 ml each of 0.2, 0.3, 0.5, and 1.0 M KCl in TM buffer.

**Gel Mobility Shift Assay**

Protein-DNA complexes were resolved essentially as described by Carthew et al. (22). The standard reaction contained 22 mM HEPES (pH 7.6), 1 mM MgCl₂, 62 mM KCl, 5 mM DTT, and 2% glycerol, with varying amounts of both poly(dl-dc)-poly(dl-dc) and either unfragmented nuclear protein extract or protein fractions from chromatography. The mixture was incubated at 100 °C for 1 min, followed by the addition of between 0.5 and 2 ng of end-labeled IAP LTR fragment. When competitor DNA fragments were present in the reaction, they were added prior to the first 20 min of incubation. After incubating for an additional 20 min at 30 °C, the entire 30-μl reaction was loaded onto a polyacrylamide gel (acylamido- bisacrylamide = 40:1) of 4% which had been pre-electrophoresed for 2 h at 30 mA. Gels were run in either 10 mM Tris-Cl (pH 7.5), 3% polyacrylamide, 1 mM EDTA, or 45 mM Tris-borate, 45 mM borax, 1% acetic acid, 1 mM EDTA. The gels were electrophoresed at 4 °C with recirculated buffer. Following electrophoresis, the gel was fixed in 5% trichloroacetic acid for 30 min, dried, and exposed to x-ray film.

**DNAse I Protection Assay**

The DNase I footprinting procedure was carried out as described (23). Each reaction contained 2 ng of a 5'-end-labeled fragment that includes the entire LTR from IAP 8IC (370 bp). Regions of DNA that were protected from DNase I cleavage were identified by chemical cleavage sequencing reactions of the same DNA fragments (24).

**Southwestern Blotting Assay**

SDS-polyacrylamide gel electrophoresis was performed as described by Leemlli (25). Nuclear extracts or column fractions were separated on an 8.75, 12, or 17% SDS-polyacrylamide gel, under reducing conditions, denatured by shaking gently in 50 mM NaCl, 10 mM HEPES (pH 7.9), 20 mM EDTA, 0.1 mM DTT, and Triton X-100 (at a ratio of 3:1 by weight to SDS) for 2 h, and electrophoretically transferred to a polyvinylidene difluoride membrane (Immobilon-P, Millipore Corporation). The membrane was stained with Ponceau S to confirm that sufficient protein was transferred, and the gel was stained with Coomasie Blue. The membrane was soaked in a solution of Southwestern buffer (50% non-fat dry milk (BLOTTO), 50 mM NaCl, 10 mM HEPES (pH 7.9) and 1 mM EDTA) for several hours at room temperature. A randomly primed oligonucleotide (>1 × 10⁶ cpm/μg), corresponding to a sequence from the IAP 8IC LTR (15), was added to 30 ml of Southwestern buffer. Binding of the oligonucleotide to the filter-bound proteins was carried out at room temperature for 3 h with gentle shaking. This was followed by four 30-min washes with Southwestern buffer and a 1-h rinse with 100 mM NaCl, 10 mM HEPES (pH 7.9), and 1 mM EDTA. The membrane was air dried and autoradiographed with an intensifying screen (26, 27).

**CAT Enzyme Assays**

The vector used for transfections (pCAP) and its construction has been previously described (15). Cell extracts were isolated 48 h after transfection, and assays for CAT enzyme activity were performed as described by Gorman et al. (28), with 0.25 μCi of [³²P]chloramphenicol (New England Nuclear Corp.) as substrate. Assays were performed on cell extracts using equal number of cells titrated to fall within the linear range of conversion of substrate to its acetylated forms. The percentage of conversion was determined by cutting out the appropriate region from the thin layer chromatography plate and counting in a toluene-based scintillation fluid.

**DNA Fragments and Oligonucleotides**

The IAP LTR 81 fragment (370 bp) was isolated from a recombinant version of the plasmid M13 mp18, which contained the IAP LTR 81 inserted into the XbaI site of the M13 mp18 polynucleotid region, by digesting with EcoRI and either Suer or Smal. The IAP LTR 81 fragment (500 bp) was also isolated from a recombinant version of M13 mp18 by digestion with EcoRI and either HanHl or HindIII.

Two methods were used to label DNA fragments used in the mobility shift assays and DNase I protection experiments. In the first, recombinant vectors were linearized with EcoRI, and the 3' ends were labeled with [³²P]dADP and the Klenow fragment of DNA polymerase I. The probe was then isolated by cutting with the second restriction enzyme and separating the fragments by gel electrophoresis. Alternatively, the LTR fragments were cut from the vectors using both restriction enzymes and purified by gel electrophoresis. The 3' ends of the LTR fragments were then labeled with [³²P]dADP and the Klenow fragment of DNA polymerase I. Oligonucleotides were synthesized on an Applied Biosystems DNA synthesizer and purified by either gel electrophoresis or high performance liquid chromatography. Each oligonucleotide containing one strand was synthesized and then annealed (29). The sequences of the IAP LTR 81 oligonucleotides are shown in Fig. 5. Synthetic
oligonucleotides were labeled at the 5' ends with T4 polynucleotide kinase and \( \gamma^{32}P \)ATP.

**RESULTS**

Deletion Analysis of the IAP 81 LTR—To determine which regions of the IAP LTR are involved in regulating transcription, several deletion mutant plasmids were constructed. Each contained the bacterial CAT gene inserted 3' to either the entire IAP 81 LTR, LTR 81 with the 5' 47 bp deleted, or only the 5' 47 bp of LTR 81 fused to the bacterial CAT gene (Fig. 1, A and B). These plasmids were transfected into PCC3 cells, and the relative promoter activity of each construct was assessed by comparing the amount of CAT protein produced in each case. IAP LTR 81 was an efficient promoter of the CAT gene (Fig. 1A, lane 1), while the construct containing LTR 81 with the 5' 47 base pairs deleted produced a level of CAT gene transcription only 20% of that found with the entire LTR 81 (Fig. 1A, lane 3). Analysis of the first 47 base pairs of LTR 81 showed it alone was unable to promote CAT gene transcription (Fig. 1A, lane 2). These results indicate that sequences within the first 47 bases (Fig. 1B, −212 to −160) of the LTR are responsible for the enhancement of transcription above the basal level as was previously reported (15).

Footprint Analysis of Protein-binding Sites within LTR 81—

![Footprint Analysis of Protein-binding Sites within LTR 81](image)

**Fig. 1.** Assay of CAT activity in PCC3 cell extracts and nucleotide sequence of the IAP LTR 81C U3 region. Panel A, PCC3 cells (7.5 \( \times \) 10⁵) were transfected with 20 \( \mu \)g of a CAT construct. Extracts (100 \( \mu \)l) were prepared 48 h after transfection, and 20 \( \mu \)l of each extract was incubated for 30 min at 37°C. The entire LTR adjacent to the CAT gene (lane 1). Only the 5' 47 base pairs of LTR 81 adjacent to the CAT gene (lane 2). LTR 81 minus the 5' 47 base pairs adjacent to the CAT gene (lane 3). The CAT gene and vector only (lane 4). Panel B, the nucleotide sequence of the U3 region of IAP LTR 81C is shown. The numbers are the distances in base pairs from the transcription start site (15, 38). The boundary of the U3 and R region is noted. A number of transcription regulatory sequences are identified: a glucocorticoid-responsive element (GRE) (39, 40), the SV40 core enhancer element (enh. core) (15, 41), an activation transcription factor/cAMP-responsive element (ATF/CRE) (42), a CCAAT box (CAT box) (43), a CCAAT/enhancer-binding protein site (C/EBP) (23, 44), and a TATA box (TATA) (45). The arrow (↓) (5' 47) indicates the position where PvuII cut to form the two deleted fragments used in the CAT transfections.

DNase I protection assays were used to determine the regions of the LTR to which nuclear proteins bind and, specifically, if any nuclear factors interact in the 5' 47-bp region. Results show PCC3 nuclear extracts protected IAP LTR 81 fragments in three regions (Fig. 2). One footprint spanned the nucleotides −168 to −211 of LTR 81. The other protected regions covered bases −19 to −69 and −72 to −97 of LTR 81. The latter two regions contain a number of promoter and proximal promoter elements, including a TATA box, a CCAAT box, and two sequences highly similar to the sites recognized by the CREB, AP1, and C/EBP families of proteins (Fig. 1B).

Fractionation of Nuclear Proteins That Interact with Different Regions of IAP LTRs—When the PCC3 nuclear extract was incubated with a radiolabeled DNA fragment containing the LTR from either IAP 81 or IAP 14, two protein-DNA complexes were formed (Fig. 3A, compare lanes 2 and 3). They have been identified as complex 1 and complex 2. Whether LTR 81 or LTR 14 was used, complex 1 (and, similarly, complex 2) appeared to be the same because, first, the mobility of the protein-DNA complexes are identical and, second, when fragments from LTR 81 were used as nonradioactive competitors against labeled LTR 14, the formation of radioactive LTR 14-protein complexes was selectively prevented (data not shown). The protein-DNA complexes were specific because an excess of unlabeled LTR 14 successfully competed for binding of the PCC3 proteins to the radioactive LTR 14 probe (Fig. 3A, lane 4), but a fragment of pBR322 DNA was unable to prevent the formation of the complexes (Fig. 3A, lane 5). In addition, bovine serum albumin failed to bind to the LTR 14 probe (Fig. 3A, lane 6).

To understand the processes involved in transcription regulation, it is necessary to have as many as possible of the components of the macromolecular machinery isolated and identified. Toward this end, the PCC3 nuclear extract was fractionated by DEAE column chromatography. Complex 2 activity did not bind to the DEAE column in 0.1 M KCl and was recovered in the flow through and wash fractions. Complex 1 activity, however, was retained by the column and was eluted with 0.2 M KCl. The fractions containing complex 2 activity were pooled, concentrated, and further fractionated on a heparin-agarose column. An additional LTR 81 binding activity (identified as complex 3) was detected in the 0.2 M
poly(dI-dC)-poly(dI-dC) with and without competitor DNA, and incubated at 30 °C for 20 min. The entire LTR, end-labeled with 32P-labeled LTR 81 and 2-4 pg of poly(dI-dC)-poly(dI-dC). Column fraction numbers and the concentration of KCl necessary to elute each fraction are also indicated. Panel C, PCC3 nuclear extract and DEAE fractionated nuclear proteins of 46,000 and 28,000 daltons (Fig. 5A, lane 1). When oligo(A) was used to reprobe the same stripped filter, only the 28,000-dalton species was the major protein to interact with the extreme end of U3. Binding of oligo(A) to the 46,000-dalton band was not observed (Fig. 5A, lane 2). Instead, the 46,000-dalton protein binds to oligo(D) (Fig. 5C). This corresponds to the interaction forming complex 1 (Fig. 4A).

To further investigate whether the 28,000-dalton protein is responsible for complex 3, proteins retained on the heparin-agarose column were subjected to a similar Southwestern analysis (Fig. 5B). Although the 0.2 M KCl eluate from the heparin column still contains many proteins (data not shown), the only one to bind the oligonucleotide probes was the 28,000-dalton protein. With comparable specific activities and equal numbers of moles among the probes, the oligo(B) probe bound less than oligo(A) and oligo(C), reinforcing the importance of having nucleotides -210 to -195 present to maximize this interaction. In addition, oligonucleotides A, B and C bound to the 28,000-dalton protein in heparin-agarose column fractions containing complex 3 (compare Figs. 3C and Fig. 5B, searched for another regulatory protein(s) that interacts with this extreme end (-184 to -211) of the U3 region, to determine whether it may be the protein responsible for complex 3.

Identification of 5' U3-Binding Proteins and Protein Involved in Complex Three Formation—Proteins from PCC3 nuclear extracts as well as from the pooled heparin 0.2 M KCl fractions were separated by SDS-polyacrylamide gel electrophoresis, immobilized onto filters, and probed with a set of four oligonucleotide probes corresponding to the 5' end of IAP LTR 81 (Fig. 5D). Oligo(A) (-210 to -195) and oligo(B) (-199 to -188) cover two overlapping regions at the 5' end of the U3 region, while oligo(D) spans base pairs -183 to -168. All three are parts of oligo(C) (-210 to -168) that covers the region shown to be essential for IAP LTR-directed transcription. Oligo(C) binds predominantly two nuclear extract proteins of 46,000 and 28,000 daltons (Fig. 5A, lane 1).

Fig. 4. Comparison of DNase I footprinting of LTR 81 by PCC3 nuclear extract and DEAE fractionated nuclear proteins. Fraction 15 of the DEAE-cellulose column (A, DEAE 0.2 M) or fraction 6 of the DEAE-cellulose column (B, DEAE 0.1 M) were used to footprint LTR 81 and compared with those by the unfracionated nuclear extract (NE). Conditions for the reactions were as described in Fig. 2.

KCl fractions eluting from the heparin column (Fig. 3C).
the electrophoresis on an 8.75% gel, transferred, and probed with oligo(C) (lane 1). The same filter, after stripping and reprobed with oligo(A) (lane 2). Panel B, fractions from the heparin-agarose column that showed complex 3 activity (fractions 14, 16, 18) or did not (fraction 9) in the mobility shift assay were used (see Fig. 3C). Proteins were separated on a 17% gel and probed with oligo(A) (lanes 1–3), oligo(B) (lanes 4–6) or oligo(C) (lanes 7–9). Markers, lane M. Panel C, 20 µg of either PCC3 (lane 1) or LTK− (lane 2) nuclear extract probed with oligo(D). Panel D, the sequences of the four synthetic oligonucleotides (A–D) used in this study. See “Materials and Methods” for detailed descriptions of procedures.

**FIG. 5. Identification of LTR 81-binding proteins by Southwestern analysis.** Panel A, 50 µg of PCC3 nuclear extract was subjected to electrophoresis on an 8.75% gel, transferred, and probed with oligo(C) (lane 1). The same filter, after stripping and reprobed with oligo(A) (lane 2). Panel B, fractions from the heparin-agarose column that showed either complex 3 activity (fractions 14, 16, 18) or did not (fraction 9) in the mobility shift assay were used (see Fig. 3C). Proteins were separated on a 17% gel and probed with oligo(A) (lanes 1–3), oligo(B) (lanes 4–6) or oligo(C) (lanes 7–9). Markers, lane M. Panel C, 20 µg of either PCC3 (lane 1) or LTK− (lane 2) nuclear extract probed with oligo(D). Panel D, the sequences of the four synthetic oligonucleotides (A–D) used in this study. See “Materials and Methods” for detailed descriptions of procedures.

**FIG. 6. Southwestern analysis of nuclear extracts from various cell types.** Equal volumes of different nuclear extracts were electrophoresed and separated on a 17% SDS-polyacrylamide gel and transferred to a polyvinylidene difluoride membrane. Panel A, proteins on the membrane were probed with radioactive oligo(A). Panel B, proteins on the gel were stained with Coomassie Blue dye. The source and amount of protein, undifferentiated PCC3 (PCC3, 24 µg), HeLa (HeLa 18 µg), mouse fibroblast (LTK− 12 µg), MOPC-315 myeloma (MOPC, 31 µg), and differentiated PCC3 (dPCC3, 31 µg). M, Markers.

Fig. 6. Southwestern analysis of nuclear extracts from various cell types. Equal volumes of different nuclear extracts were electrophoresed and separated on a 17% SDS-polyacrylamide gel and transferred to a polyvinylidene difluoride membrane. Panel A, proteins on the membrane were probed with radioactive oligo(A). Panel B, proteins on the gel were stained with Coomassie Blue dye. The source and amount of protein, undifferentiated PCC3 (PCC3, 24 µg), HeLa (HeLa 18 µg), mouse fibroblast (LTK− 12 µg), MOPC-315 myeloma (MOPC, 31 µg), and differentiated PCC3 (dPCC3, 31 µg). M, Markers.

fractions 14, 15, 16 and 18), but not in a fraction missing the complex 3 activity (compare Figs. 3C and Fig. 5B, fraction 9). Interactions between these proteins and oligonucleotides are specific because an oligonucleotide containing a mutated C/EBP-binding site failed to interact with any proteins on the filters (data not shown). In addition, the LTR 81 oligonucleotides were unable to bind to many other proteins from the fractions and the proteins used as molecular weight markers, both present on the filters.

Enrichment of the 28,000-Dalton Protein in Mouse Cells Having High IAP Activity—To determine the relationship of this 28,000-dalton protein to IAP transcription, nuclear extracts were prepared from five cell types and tested by the Southwestern assay with oligonucleotide A. Only extracts from undifferentiated PCC3 cells and MOPC-315 myeloma cells contained detectable amounts of the oligonucleotide A-binding protein. Binding was greatly reduced in HeLa, LTK−, and differentiated PCC3 cells, when comparable amounts of nuclear extract were used (Fig. 6). MOPC-315 and PCC3 cell lines are known to produce IAP transcripts in large amounts (14, 18). HeLa cells do not make IAPs. LTK− and differentiated PCC3 cells do so in markedly reduced amounts (19, 36).

**DISCUSSION**

The embryonal carcinoma cell line PCC3-A/1 is an excellent system for the study of the mechanisms that regulate IAP gene transcription. Undifferentiated PCC3 cells actively transcribe endogenous IAP genes (11, 12, 18) and express IAP-CAT chimeric constructs in a transient transfection assay (15, 17). Nuclear extracts from PCC3 cells contain the components necessary for in vitro transcription from IAP LTR-promoted templates (data not shown). This current study has shown that PCC3 nuclei contain a set of proteins that bind within the IAP LTR at specific DNA sequences that are either putative or proven cis-acting transcription regulatory elements.

Two regions of LTR 81 bound proteins from PCC3 nuclear extracts, as determined by DNase I protection assays. These two regions are coincident with sequences of LTR 81 that are required for full transcriptional activity to occur (15). A set of protein-binding sites are located within the 100 bp immediately 5′ of the transcription initiation site. This region is critical for basal transcription of the CAT reporter gene, promoted by LTR 81, when transfected into PCC3 cells (15 and Fig. 1). Other investigators (16) using the MIA14 IAP LTR and nuclear extracts from MOPC-315 cells and a human kidney carcinoma cell line, 293, observed DNase I-protected regions in the proximal promoter region with the identical 5′ terminus as ours and spanning to within four base pairs of the end of their fragment. (The MIA14 LTR fragment terminated at a HaeIII site equivalent to position −73 in LTR 81.) Each of the footprints we observed in this proximal region has at least some similarity to the binding site of a purified eukaryotic transcription factor. Within the footprint spanning −72 to −97, there is an ATF/CRE site, a CCAAT box, and a sequence containing the core of another ATF/CRE palindrome, ACGTC. It has yet to be proven that any of the cognate transcription factors bind to these sites within the IAP LTR and participate in regulating transcription. We have, however, recently discovered that a recombinant version
of C/EBP can bind to the two ATF/CRE sites and the CCAAT site. Binding to the ATF/CRE is in agreement with two other reports that have investigated the role of C/EBP in transcription regulation (30, 31).

The second region within LTR 81 that bound undifferentiated PCC3 nuclear proteins is a stretch of 43 base pairs at the 5′ terminus (−168 to −211). This portion of the LTR is very important for enhancing transcription, as deletion of this region in the LTR-promoted CAT transfections severely reduced the level of CAT activity. The binding of PCC3 proteins to this region, as assayed by DNase I protection, has been corroborated by others using similar IAP LTRs and nuclear extracts from other cell types (16). When PCC3 nuclear proteins were passed through a DEAE-cellulose column, 0.2 mM KCl fractions contained a DNA binding activity that protected bases −183 to −168 on IAP LTR 81. These same fractions could form only a single protein-DNA complex (complex 1) with LTR 81 in a gel mobility shift experiment. The protected 16-base pair sequence is not similar to any known transcription factor binding sites, nor does it have any distinguishing sequence features, such as regions of symmetry. However, this sequence is well conserved among all IAP genes that have been analyzed. Located within this conserved sequence is a HhaI site containing a methylation-sensitive cytidine/guanidine pair (position −173). Falzon and Kuff (32) have shown that methylation of this site in the 5′ LTR from the IAP genomic clone MIA14, caused a 2-fold reduction of CAT activity in 293 cells. We have recently found that hypomethylation of this site correlates with expression of the MIPP gene in mouse placenta. Using Southwestern blotting, we have identified a 46,000-dalton protein in PCC3 cells that binds only to this 16-base pair sequence. We are purifying this 46,000-dalton protein and studying its role in regulating transcription of IAP LTR-driven genes.

The LTR 81 sequence between −184 and −211 appears to bind a 28,000-dalton nuclear protein in undifferentiated PCC3 cells. Between −195 and −209, there are two potential binding sites arranged with dyad symmetry around a single base pair at position −202. A single half-site has the sequence 5′TGCGPuGC3′. Such an arrangement for the binding sites of transcription regulatory proteins is well documented (30, 33–35), and proteins bind to such sites as dimers. This particular sequence motif in IAP LTR 81 is well conserved among IAP LTRs and has not been described previously. Also within this region is a sequence (−205 to −197) that when inverted closely matches the current consensus binding site for Sp1 (36). In addition, this region binds the 28,000-dalton protein of nuclear extracts isolated from cell lines that actively transcribe IAP genes. The protein-DNA complex was prominent in PCC3 nuclear extracts and even more intense in extracts derived from the myeloma cell line MOPC-315. Myelomas, and MOPC-315 cells in particular, produce large amounts of IAP transcripts (14). By contrast, HeLa cells do not transcribe IAP genes and contain no detectable protein that interacts with this sequence. Especially interesting was the absence of detectable binding activity from PCC3 cells that had been induced to differentiate in culture by exposure to retinoic acid. Northern blot analysis comparing PCC3 cells before and after differentiation indicates that the steady-state level of IAP transcripts drops markedly following differentiation (18, 37). In addition, we have seen a drop in the binding affinity for LTR 81 by nuclear proteins derived from differentiated PCC3 cells. These results suggest that the 28,000-dalton protein we have identified by Southwestern analysis may be involved in IAP-specific transcription regulation.

The presence of an inverted Sp1 consensus sequence raises the possibility that Sp1 and our 28,000-dalton protein are related. The 28,000-dalton protein could be a proteolytic fragment of the approximately 100,000-dalton Sp1 protein. This is unlikely because HeLa cell nuclear extract, the richest source of Sp1, does not contain the 28,000-dalton, IAP LTR-binding protein. It is also possible that Sp1 and the 28,000-dalton protein bind to the same site or overlapping sites. We observed that HeLa cell nuclear extract exhibited weak binding to oligo(A) at a position corresponding to a 100,000-dalton protein (size of Sp1) on a SDS-polyacrylamide gel (Fig. 6A). A similar finding was observed on another Southwestern analysis, when oligo(A) was used to probe a filter containing PCC3 nuclear proteins (Fig. 5A, lane 2). Determining whether any relationship exists between the 28,000-dalton protein and Sp1 will have to wait until this protein is purified and sequenced. This work is in progress.

These studies show that the regulation of transcription from IAP LTRs very likely involves concerted interactions between groups of DNA control elements and sequence-specific DNA-binding proteins. A key set of novel elements that markedly stimulates transcription has been identified along with their cognate binding proteins. Studies are in progress to purify these proteins to homogeneity, making it possible to examine how the individual proteins interact with the IAP LTR DNA and, very likely, with one another.

Acknowledgment—We thank Dr. David E. Mold for advice and critical reading of the manuscript and for conducting Southwestern analysis with oligo(D).

REFERENCES

2 M. Zierler, unpublished results.
3 D. E. Mold and R. C. C. Huang, manuscript in preparation.