Multiple protein binding sites within the ovalbumin gene 5'-flanking region: isolation and characterization of sequence-specific binding proteins

Martine Pastorcic, Milan K.Bagchi, Sophia Y.Tsai, Ming-Jer Tsai and Bert W.O'Malley*

Department of Cell Biology, Baylor College of Medicine, Houston, TX 77030, USA

Received April 10, 1989; Revised and Accepted June 30, 1989

ABSTRACT

To understand the molecular basis of the steroid hormone regulated expression of the ovalbumin gene, we sought to identify and isolate nuclear factors from chicken oviduct which interact specifically with the ovalbumin promoter. Using DNase I footprinting and mobility shift assays, we have defined at least four distinct protein binding sites, OV-150, OV-220, OV-250 and OV-330, in the promoter region between -100 to -400. Binding competition and protein fractionation studies revealed the existence of two distinct proteins, each recognizing two promoter sites: Both OV-330 and OV-250 are recognized by one protein factor which is distinct from the one binding to both OV-220 and OV-150. The location of the DNase I footprints coincides with those of in vivo chromatin hypersensitive sites (1). The OV-330 site is located in a sequence area required for the repression of the gene in the absence of hormone (2). The factor binding to OV-330 has been substantially purified and renaturation experiments indicate that the binding activity is associated with a polypeptide(s) of Mr 40K.

INTRODUCTION

The tissue specific expression of the ovalbumin gene and its induction by steroid hormones (3,4) make it an attractive model to study gene expression. In the 5'-flanking region of the gene, distinct promoter elements such as the TATA box (5,6) and the COUP box (7,8,9,10) have been characterized. These elements are required for the basal level of expression of this gene in heterologous cells such as HeLa (7) and for the initiation of in vitro transcription in HeLa extracts (5,8,9). The COUP transcription factor which recognizes the -70 to -90 region of the ovalbumin promoter has been purified from both HeLa and oviduct cells (11,12). Further upstream around -1100, we have identified recently a hormone independent enhancer element which binds an NFI-related nuclear factor present in oviduct cells (13). However, the basis of the tissue specific

expression of the ovalbumin gene remains to be elucidated. Recently, Dierich et al. (14) and Gaub et al. (2) microinjected a series of ovalbumin promoter mutants fused to the SV40 T antigen gene into oviduct primary culture cells and assayed their expression by immunofluorescence. They delineated sequences that are crucial for ovalbumin gene expression and steroid hormone induction in the first 400 base pairs upstream from the initiation site of transcription. Of particular interest was the observation that in the absence of steroid hormone, the gene was under negative regulation which required a *cis* element located between -426 and -296 (2); the repression appeared to be released by steroid hormones.

The understanding of the mechanism of expression of the ovalbumin gene in oviduct cells will require a detailed mutagenic analysis of the promoter. An alternative and complementary approach will be the functional reconstitution *in vitro* of the steroid-regulated transcription of the gene starting from individually purified components. In this report we describe the identification and isolation of protein factors from chick oviduct nuclear extract which interact specifically with discrete ovalbumin promoter sequences between -100 and -426. We have examined their interrelationships by using binding competition and protein fractionation studies.

EXPERIMENTAL PROCEDURES

DNA mobility shift assays - Protein fractions were incubated with 2 to 5 fmol of labeled DNA probe in 12 mM Hepes (pH 7.9), 50 mM NaCl, 0.5 mM DTT, 0.1 mM EDTA and 12% glycerol. Poly dI-dC (Pharmacia) was included in the reaction as a nonspecific competitor (15). Gel electrophoresis of the binding mixture was as described previously (9).

DNA probes - A 5'-deletion construct of the ovalbumin promoter pSV₄OG (-450) (8) was end-labeled with the Klenow fragment of DNA polymerase and [α^{32} P]dCTP at the Cla I site (position -450) and digested with Ava II (position -135). The resulting fragment then was purified by gel electrophoresis. A 3'-deletion construct (pOV-3' Δ -191) similarly was used to prepare a probe labeled in the opposite strand. The latter fragment contained sequences from the endpoint of the deletion at -191 (labeled at the Cla I site) to -426 (Rsa I site). Another deletion mutant pOV-3' Δ -103 was used to prepare a probe containing sequences from -103 to -269 (AluI).

Oligonucleotides containing portions of the ovalbumin promoter were synthesized with added Bam HI ends. They were end labeled as described above with [α^{32}]dCTP. Oligo-330 contained ovalbumin sequences from -340 to -304;

oligo-280 from -296 to -261, and oligo-140 from -166 to -133. Another oligonucleotide contained sequences -215 to -255. The COUP oligonucleotide (-92 to -66) was described previously (9). The specific activity of the DNA probes was usually 1 to 2 x 10⁷ cpm/pmol of fragment.

DNase I footprinting - DNA binding reactions (20 µl) were carried out as described for DNA mobility shift assays, except that incubations were for 20 minutes at 22°C. DNase I and MgCl₂ (5 mM) were added then and incubated for another 30 seconds. The reaction was stopped by addition of 100 µl of 50 mM Tris (pH 7.5), 50mM NaCl, 10 mM EDTA, 0.5% SDS, containing 0.1 mg/ml proteinase K, and 1 µg of pBR322 as carrier, and incubated at 37°C for 15 minutes. The samples were extracted with phenol-chloroform, ethanol precipitated and loaded on 8% polyacrylamide - 7M urea sequencing gels.

Purification - All steps were carried out at 4°C. Nuclear extract obtained from chick oviduct was processed through DEAE-sephadex, phosphocellulose and heparin sepharose column chromatography steps described previously (12).

Purification of the OV-330 binding protein: Fractions from the heparin sepharose column containing the DNA binding activity specific for the -330 region of the ovalbumin promoter were pooled, diluted to 100 mM KCl by addition of buffer D and loaded onto a 3 ml double strand DNA cellulose column (Pharmacia) preequilibrated with buffer D plus 100 mM NaCl. The activity was eluted with a 30 ml linear gradient from 100 to 500 mM NaCl and appeared at 300 mM NaCl. The protein finally was purified by two consecutive DNA affinity steps.

DNA affinity chromatography - The preparation of the specific DNA column using concatemers of OV-330 was essentially as described previously (11,12,16). About 5µg of oligonucleotide was bound per 100 µl of resin. Protein fractions from the calf thymus DNA cellulose column (4 mg) were first concentrated on a 0.7 ml phosphocellulose column and step-eluted with 0.8 ml of 500 mM NaCl in buffer D. For the first DNA affinity step, the protein (3.5 mg) was incubated on ice for 60 minutes with 1 ml of resin and 3.5 mg of poly dI-dC in buffer D containing 60 mM NaCl in a total volume of 5 ml, with gentle agitation to keep the resin in suspension. The column then was packed quickly and washed extensively with 30 column volumes of buffer D containing 100 mM NaCl. Elution was carried out in a stepwise manner with 1.5 ml of 150 mM NaCl, 400 µl of 200 mM NaCl, 1 ml of 300 mM NaCl and 1 ml of 400 mM NaCl and finally with 1 ml of 1M NaCl in buffer D. 160 µl fractions were collected into siliconized tubes. The active fractions were pooled and applied to a second DNA affinity column as follows. The pooled fractions (1.5 ml) were incubated with 200 µg poly dI-dC and 200 µl of resin in 10 ml of buffer D containing 50 mM NaCl

Figure 1: DNase I footprinting analysis of the chicken ovalbumin promoter with oviduct protein fractions. (A) DNase I protection on the top strand with 1 ng of DNA fragment from -103 to -269 and 400 ng poly dI-dC. Lanes marked (+) indicate incubations including 1 μ g of the heparin sepharose pool (0.25 to 0.45 M KCl). In lanes marked (-) no protein was added. The footprint is indicated by brackets and an arrow shows the position of the hypersensitive site at -152. **(B)** 0.4 ng of DNA fragment from -191 to -426 labeled at the 3' end on the top strand. 400 ng of poly dI-dC was included with 2 μ g of the heparin sepharose pool. The footprints observed: OV-220, OV-250, OV-330 are indicated by brackets. The arrows mark the hypersensitive sites at -235 and -236. **(C)** The heparin sepharose pool was fractionated further through a DNA affinity column containing concatemers of sequences from -215 to -255. Proteins absorbed on the column were eluted with 0.5 M KCl. Footprinting assays with this eluted fraction (2 μ l) included 0.4 ng of fragment from -191 to -426 and 100 ng poly dI-dC.

and 50 ng/ml BSA for 60 minutes at 4°C. The column was packed and washed as described for the first step. The column was eluted with 300 μ l of 150 mM NaCl, 300 μ l of 200 mM NaCl, 300 μ l of 300 mM NaCl, 400 μ l of 400 mM NaCl and a final washing was performed with 1 ml 1M NaCl in buffer D. 70 μ l fractions were collected. All fractions were frozen quickly in a dry ice - ethanol bath and stored at -70°C.

Figure 2: Competition of protein binding to the -215 to -255 region by mobility shift assay. 0.2 ng of radiolabeled oligonucleotide containing sequences from -215 to -255 was incubated at 22° for 20 minutes with the affinity column fraction used in the experiment described in Figure 1(C). Assays contained 100 ng poly dI-dC and various amounts of unlabeled competitor. (A) Lanes 1, 4 and 7: no specific competitor added. Lanes 2 and 3, 5 and 6 and 8 and 9: 20 and 40 fold molar excess of unlabeled oligonucleotides -215 to -255, oligo-330 and COUP (-92 to -66), respectively. (B) Lanes 1 and 5: no specific competitor added. Lanes 2 to 4 and 6 to 8 contained 20, 30 and 40 fold molar excess of oligonucleotides -215 to -255 and oligo-150, respectively.

Renaturation of protein after SDS-polyacrylamide gel electrophoresis

Renaturation was performed according to the procedure of Hager and Burgess (17) modified as follows. A 160 μ l sample of fraction 30 from the first DNA affinity column was fractionated through a 15% polyacrylamide gel containing SDS. The electrophoresis was performed at 25 mA at 4°C. A part of the gel was cut into slices of various molecular weight range (Figure 5A). The other part, where the position of the slices was marked by notches, was stained by silver (18). Each slice was placed into a dialysis bag containing 250 μ l of Laemmli running buffer with 0.5 mg/ml BSA, and proteins were electroeluted at 10 volts/cm for 3 hours at 4°C. The content of the dialysis bags was collected into siliconized tubes and precipitated by addition of 4 volumes of acetone at -20°C. Three additional precipitations by acetone ensured the removal of SDS.

Finally, the dried pellets were dissolved into 200 μ l of 6 M guanidine hydrochloride and incubated at room temperature for 30 minutes, and then dialyzed for 14 hours at 4°C against buffer D [20 mM Hepes (pH 7.9), 0.5 mM DTT, 0.1 mM EDTA, 20% glycerol] plus 100 mM NaCl. Samples were assayed immediately for binding activity or frozen in a dry ice - ethanol bath and stored at -80°C.

RESULTS

Detection of Protein Factors Recognizing the Ovalbumin Promoter by DNase I Footprinting

To identify DNA binding proteins from oviduct nuclear extracts which recognize specifically the ovalbumin promoter, we end-labeled various fragments of the promoter from -103 to -426 and used them as probes in DNase I footprinting and mobility shift assays. We analyzed crude nuclear extracts and fractions from various steps of the purification of the COUP factor (12). After the DEAE and phosphocellulose columns, fractions containing the COUP factor were loaded onto a heparin sepharose column. The bound proteins were eluted with a linear gradient of 0.2 to 0.6 M KCl. The COUP factor appeared at 0.5 M KCl in a peak distinct from the majority of DNA binding proteins which were eluted between 0.25 and 0.45 M KCl; these early fractions were pooled and tested for specific DNA binding activity. When the heparin sepharose pool was incubated with a DNA probe including sequences from -103 to -269, one area of protection from DNase I appeared from -142 to -162 which included a hypersensitive cut at -152 (Fig. 1A). Footprinting assays performed using a probe containing sequences from -191 to -426 revealed several areas of protection (Fig. 1B). The most upstream footprint was located around -330. A small but discernible footprint was centered around -280. A strong area of protection spanned -217 to -268 and included two hypersensitive sites at position -234 and -235. The size of this footprint suggested that it might represent the binding of more than one protein to adjacent sequences. To clarify this point we further fractionated the heparin sepharose 0.25 to 0.45 M KCl pool on a DNA affinity column containing concatemers of the sequence -215 to -255. The proteins were applied to the column in the presence of nonspecific competitor (poly dI-dC) and eluted in a single step with 0.5 M KCl. DNase I footprinting assays carried out with the 0.5 M KCl eluate revealed a predominant area of protection from -217 to -238 only (Fig. 1C). Thus, the large footprint extending from -217 to -268 as originally observed with the heparin sepharose pool seemed to be composed of the juxtaposed binding sites of two distinct proteins. The specific-DNA column efficiently retained a protein recognizing the proximal site in the footprint -217 to -238 (OV-220); it did not retain the protein conferring protection to

Figure 3: Detection of a specific DNA-protein complex with oligo-330 in oviduct nuclear extract. 5 µg of nuclear extract was incubated with 4 fmol of end-labeled oligo-330 in 20 µl binding mixtures. In lanes 1 to 5 the assays were performed in the presence of increasing amounts of poly dI-dC. In lanes 7 to 16 the reactions contained 2 µg of poly dI-dC plus various amounts of unlabeled oligonucleotide competitors (oligo-330, oligo-280 or COUP) as indicated below the lanes. No protein was added in lane 6. The position of the specific complex is marked by an arrow.

the region upstream of -238. This interpretation was supported further by the subsequent purification of the protein recognizing the -238 to -268 region (OV-250) as described below.

Binding Competition Studies

To analyze the relationships between these binding proteins, we performed a series of competition experiments using mobility shift assays. We synthesized double-stranded oligonucleotides containing the putative binding sites defined by DNase I footprinting. The oligonucleotide sequences spanned -133 to -166, -215 to -255, -304 to -340 and included binding sites OV-150, OV-220 and OV-330, respectively. These oligonucleotides were end-labeled and used as probes in DNA mobility shift assays. Each of the oligonucleotides formed a stable retarded complex in the presence of the nonspecific competitor poly dI-dC and the appropriate protein fraction indicating that the protein factors did indeed recognize these sequences (data not shown). The specificity of these complexes was analyzed further by competition experiments. Using radiolabeled oligonucleotide -215 to -255 as a probe, a complex was formed with the DNA affinity purified OV-220-binding protein fraction; this complex was efficiently

Figure 4: Purification of the OV-330 binding protein by DNA affinity chromatography. (A) Binding assays. 20 μ l reactions contained 4 fmol of 32 P-labeled oligo-330, 50 ng of poly d(I-C), 0.3 μ l of the DNA affinity II fraction as indicated below the lanes and 3 μ g BSA. Incubation was on ice for 30 minutes. The position of the specific complex is indicated by an arrow. **(B) SDS polyacrylamide gel analysis.** 40 μ l of fractions across the activity peak (indicated above the lanes) were electrophoresed on a denaturing 15% polyacrylamide gel alongside with 60 ng of each of molecular weight standards (M) and BSA (B). The protein pattern was visualized by silver staining. Arrows indicate the size of the molecular weight standards.

competed by a 50-fold molar excess of the same unlabeled oligonucleotide (Fig. 2A). The binding was not competed by the same amount of either oligo-330 or COUP oligonucleotide, but was competed efficiently by oligo-150 (Fig. 2B). In the converse experiment, the formation of the complex with labeled oligo-150 was

TABLE I - PURIFICATION TABLE

	TOTAL PROTEIN (mg)	TOTAL UNITS	SPECIFIC ACTIVITY (units/mg)	PURIFICATION FOLD	YIELD
Nuclear Extract	2,800	65000	23	--	100
DEAE Sephadex	2,100	44,000	22	--	67
Phospho Cellulose	140	31,000	220	10	48
Heparin Sepharose	22	25,000	1.1×10^3	50	38
DNA Cellulose (Calf Thymus)	3.5	21,000	6×10^3	272	32
DNA Affinity I	0.003	13,000	4.3×10^6	2×10^5	20
DNA Affinity II	0.0008	7,800	9.7×10^6	4.4×10^5	12

1 unit of activity is defined as the amount of protein required to bind 1 fmol of probe in DNA mobility shift assay

competed by addition of either unlabeled oligo-150 or oligonucleotide -215 to -255 (data not shown). In summary, protein binding and competition studies suggest that distinct oviduct nuclear factors recognize the COUP element (-70 to -90), OV-150 and OV-330. However, the factors binding the OV-150 and the

sequences from -217 to -238 may be identical or related. Furthermore, the purification of the protein recognizing OV-330 (described below) suggested that it might be identical or related to the factor recognizing the -238 to -268 region.

Purification of OV-330 Factor

To purify this protein, we have used a mobility shift assay with an oligonucleotide probe (OV-330) containing the entire footprint area. We first observed that a protein-DNA complex indeed can be detected in nuclear extract (Figure 3). It appears to be specific for the -330 region since it can be competed totally by 50-fold molar excess (10 ng) of the homologous oligonucleotide but not by as much as 200-fold excess (40 ng) of an oligonucleotide corresponding to the adjacent sequence OV-280 or to the "COUP" oligonucleotide.

The initial steps of the purification up to the heparin sepharose column chromatography were identical to those used for the purification of the COUP transcription factor described previously by Bagchi *et al.* (12). Briefly, crude nuclei were prepared from diethylstilbestrol stimulated chick oviducts and proteins were extracted from the nuclei with 0.5 M NaCl. The nuclear extract was loaded onto a DEAE-sephadex column equilibrated with 50 mM ammonium sulfate. Most of the OV-330 binding activity (70%) was present in the flow-through fraction. This fraction was dialyzed to 200 mM KCl, loaded onto a phosphocellulose column and eluted with a linear gradient of 200 to 700 mM KCl. A peak of specific DNA-binding activity appeared around 350 mM KCl. After dialysis to 200 mM KCl, the proteins were loaded onto a heparin sepharose column and eluted with a linear gradient of 200 to 600 mM KCl. The OV-330 binding activity eluted at 450 mM KCl and could be separated from the COUP factor eluting from 500 to 600 mM KCl. About 50-fold purification of the OV-330 binding activity was achieved at that stage. After dilution to 100 mM KCl, the heparin sepharose fraction was loaded onto a double-strand DNA cellulose column. Only 50% of the proteins applied to the column were absorbed. A linear gradient of 100 to 500 mM NaCl eluted the bound protein with the peak of the OV-330 binding activity at 300 mM NaCl (data not shown).

The DNA affinity chromatography then was carried out in the presence of a large excess of nonspecific competitor poly dI-dC as described in the Experimental Procedures. The bound proteins were eluted in several steps with increasing ionic strength from 0.1M to 0.4M NaCl. Less than 15% of the OV-330 binding activity loaded on the first column was found in the flow-through fractions. Most of the specific DNA-binding activity eluted from the column between fraction 17 (250 mM NaCl) and fraction 38 (400 mM NaCl) (data not

shown). SDS polyacrylamide gel analysis of the proteins present in the fractions across the elution profile (fractions 18, 22, and 30) (Figure 5A) shows that most of the proteins eluted from the column ahead of fraction 22, where the peak of activity appeared (data not shown). By our estimation, less than 0.1% of the proteins loaded onto the affinity column eluted in the active fractions, thus leading to an 800-fold purification by this step alone (Table I). It should be noted, however, that calculations of specific activity may be overestimated during purification if inhibitors of the DNA-binding assay are removed.

For further purification of the OV-330 binding activity a second cycle of affinity chromatography was carried out. About 25% of the activity loaded onto the column was present in the flow-through fractions. Again the activity was eluted between 250 mM (fraction 23) and 400 mM NaCl (fraction 33) (Figure 4A). SDS gel analysis indicates that the active fractions contain two major polypeptides of 38 and 40 kilodaltons (Figure 4B); they are most abundant in fractions 25 and 28, in a lesser amount in fraction 23 and hardly detectable in fraction 33. Thus, the appearance of these two polypeptides coincides with the peak of DNA binding activity (Figure 4A). The fold purification and recovery of activity at each step of the purification protocol is described in Table I.

Renaturation of the OV-330 Binding Activity After SDS Polyacrylamide Gel

To test if the 38-40 kilodalton polypeptides indeed contain the OV-330 binding activity, we performed a renaturation experiment. A sample from fraction 30 of the first DNA affinity column was fractionated on an SDS polyacrylamide gel. The gel was cut into slices of various molecular weight range as indicated in Figure 5A. The proteins from each slice were electroeluted, renatured and assayed for specific DNA binding activity (Figure 5B). In this particular and subsequent experiments, a 195 nucleotide ovalbumin DNA fragment (-426 to -232) was used as labeled probe. The proteins renatured from slice #4 formed a predominant complex of the same mobility as the native binding protein (fraction #30) and also a minor slower moving complex. 5 to 7% of the activity loaded onto the SDS gel was recovered after renaturation.

To test the binding specificities of the renatured proteins, we performed competition experiments (Figure 5C). The complexes generated by proteins obtained from gel slice #4 could be competed specifically by a 20-fold (4 ng) molar excess of unlabeled OV-330 added to the binding reaction whereas 40-fold (8 ng) excess of COUP oligonucleotide had no effect. Thus, the specific binding activity for the OV-330 region appears to copurify with polypeptides of approximately 40 kilodalton. No attempt was made to discriminate between the various proteins appearing in that size range. The minor slower moving complex

(from slice #4) also appeared to be specifically competed by OV-330. It is possible that this complex resulted from the presence of a second binding site present on the probe which includes sequences between -426 and -232. This hypothesis is supported by the DNase I footprinting data described below.

A small fraction of the OV-330 specific binding activity was recovered from slice #3 (Fig. 5B and 5C). It migrated with a slightly slower mobility compared to the native complex and was not apparent in assays with purified native protein (fraction 30) where it may be masked by a much larger amount of the native complex. This activity appeared to be contained in a polypeptide of a size slightly larger than 40 kDa in the renaturation assays and was not detectable by silver staining of SDS-polyacrylamide gels such as the one shown in Figure 4B. This complex did not appear in higher proportion in assays with fractions prior to the DNA cellulose steps (data not shown).

Characteristics of the OV-330 Binding Activity

The fractions obtained from the second DNA-affinity column where the OV-330 binding activity was detected by mobility shift assay were tested by DNase I footprinting (Figure 6). Using a promoter fragment -426 to -232 we observed two protected areas from position -342 to -324, and -267 to -237 on the coding strand. This may indicate that the same protein, or proteins of related specificity recognize the OV-330 and OV-260 sites. This conclusion was further substantiated by the results of the competition studies which showed that the OV-330 oligonucleotide could compete for the DNase I footprint at both -330 and -260 sites (data not shown).

Figure 5: Renaturation of gel purified OV-330 binding activity. (A) SDS polyacrylamide gel analysis of the DNA affinity column fraction used in the renaturation experiment. 40 μ l of fraction 30 from the DNA affinity I which was used in the renaturation experiment was electrophoresed alongside with 40 μ l each of two other fractions 18 and 22 and 1 μ l of the DNA cellulose pool (D). The positions of the slices cut out from the gel for the renaturation experiment are indicated by numbers 1 to 7 on the right hand side of the gel. (B) 160 μ l of fraction 30 from the DNA affinity column I were loaded on a 15% SDS-polyacrylamide gel. The proteins from the gel slices (indicated in panel A) were electro-eluted, renatured and tested for binding activity. 60 μ l of the renatured protein fraction obtained from each slice (indicated below the lanes), was incubated with 20 ng of poly dI-dC and 10 fmol of end-labeled ovalbumin promoter fragment from position -232 to -426 and analyzed by mobility shift assay. 0.02 μ l of fraction 30 from DNA affinity column I was used in lane C. An arrow indicates the position of the specific complex. (C). Specificity of the DNA binding activity after renaturation. Assays were performed according to the description in the legend to panel B. 4 fmol of end-labeled probe and 40 ng of poly dI-dC were used in this particular experiment. Unlabeled oligonucleotides were added as competitors in indicated amounts. In lane a: no competitor was added. Lane b contained 8 ng of COUP oligonucleotide, lanes c and d: 4 and 8 ng of oligo-330, respectively. 4 μ l of renatured protein fractions from slices 3 and 4 and 0.02 μ l of fraction 30 from DNA affinity column I were used.

Figure 6: DNase I footprinting analysis of DNA affinity column II fractions. Footprinting on the coding strand (bottom strand) was carried out with 1 fmol of end-labeled probe including sequences from -135 to -450, 50 ng of poly dI-dC and DNA affinity II fraction 25. Lanes 2 and 3 contained 1 and 2 μ g of the affinity purified fraction. No protein was added in lane 1. A Maxam and Gilbert guanosine sequencing ladder (G) was run alongside. Footprint on the noncoding strand (top strand) was carried out with a probe including sequences from -191 to -426. In lanes 1 and 2 incubations did not contain any protein. Lanes 3, 4 and 5 contained 1, 2 and 4 μ l of fraction 25, respectively. The brackets indicate the positions of the footprints at OV-250 and OV-330.

DISCUSSION

The hormonal induction of the ovalbumin gene is mediated clearly by steroid hormone receptors (3,4). In primary chicken embryonic fibroblasts co-transfection of the ovalbumin promoter with plasmids harboring the progesterone

receptor gene led to a weak induction of the ovalbumin gene by progesterone. This induction appears to be conferred by a progesterone responsive element located between -58 and +1 (19). However, the mere presence of the receptor molecules in the cell is not sufficient to establish the tissue specific transcription of this gene (14) and the short promoter area around the TATA box (-58 to +1) is able to confer only some degree of specificity to the expression of the gene (14). To identify proteins that coregulate ovalbumin gene transcription, we have characterized nuclear factors from oviduct which recognize specifically the ovalbumin promoter from the start site of transcription to -426. Their binding sites have been defined by DNase I footprinting (Fig 7A). Using competition studies and protein fractionation we have shown that some of these sites appear to bind the same, or very closely related protein(s).

Studies by Gaub et al. (2) suggested that the -295' to -425 segment of the 5'-flanking sequence of the ovalbumin gene contains a negative regulatory element(s) which represses gene expression in the absence of steroid hormone. Since the OV-330 binding site resides in this region, this factor has been purified substantially from oviduct tissue using a combination of the classical purification techniques and affinity chromatography on specific DNA affinity column. The detection of the activity in cruder fractions (nuclear extract and DEAE) was facilitated greatly by DNA mobility shift retardation assays using a labeled oligonucleotide probe, thus minimizing the interference of nonspecific DNA binding activities. The activity copurifies through several column chromatography steps with a group of polypeptides of about 40 kilodaltons.

Although the affinity chromatography selected for specific binding to OV-330, the purified protein appears to recognize a proximal site (at -250) as efficiently. Copurification of two tightly associated proteins which interact with each other to bind to the ovalbumin promoter cannot be ruled out. However, because of the similarities in DNA sequences at both sites we feel it is likely that they are recognized by the same protein (Fig. 7B). Similarly, competition studies and protein fractionation data indicate that the same protein may recognize OV-220 and OV-150, and a remarkable homology exists between sequences from -142 to -159 and sequences from -222 to -239 (Fig. 7B and 7C). We also noticed that OV-150 and OV-220 display 50 to 70% identity with the glucocorticoid/progesterone responsive elements (GRE/PRE) I, II and III of the tyrosine aminotransferase (TAT) gene (20). A high degree of similarity (70-80%) is found also in the OV-220 site and the 5'-flanking sequences from -166 to -188 of the human growth hormone gene (21), and sequences from -188 to -209 of the chicken ovomucoid gene (22). With the exception of the TAT gene, the functional

Figure 7: Interactions of oviduct nuclear factors with the ovalbumin promoter. A, B and C. Sequence of protein binding sites and the similarities between OV-330 and OV-250, OV-220 and OV-150, D. The nucleotide position upstream from the initiation site is indicated under the line. Transcribed sequences are marked (OV). Boxes indicate protein binding sites. Hatched boxes: OV-330 factor which also binds OV-250. Full boxes: OV-220 factor also binding OV-140. The weak binding site at -280, the COUP element, the TATA box and the NFI-like enhancer binding site at -1100 are indicated also. Arrows mark the DNase I hypersensitive sites observed in footprinting assays. I_B, I_C, I_D, I_E and II are the chromatin hypersensitive sites reported by Kaye et al., (1).

significance of these sequences is not clear at present. However, the presence of these closely related sequence motifs in the 5'-flanking regions of several genes which are regulated by steroid hormones may have important implications.

To localize putative regulatory sites on the ovalbumin promoter, Kaye *et al.* (1) mapped the DNase I hypersensitive sites of the gene in oviduct nuclei. Four discrete hypersensitive regions were identified upstream from the start of transcription. One of these hypersensitive regions (region I) extended from the initiation site to approximately 300 base-pairs upstream. Within this region, at least three individual sites of hypersensitivity could be distinguished: at the TATA box (Site B), at approximately -120 (Site C), at -210 (Site D), and at -400 (Site E). These sites were all oviduct specific and hormone inducible, with the interesting exception of Site E which appeared essentially upon hormone withdrawal. Figure 7D summarizes the locations of the hypersensitive sites found within the promoter region considered in this report. The sites of nuclease hypersensitivity in transcriptionally active genes often mark the sites of interactions of *trans*-acting regulatory molecules with specific DNA sequences. A comparison of the locations of our *in vitro* footprints with those of the *in vivo* nuclear hypersensitive sites is presented in Fig. 7D. The sites C and D are at or near the footprints around -150 and -220; the footprint at -330 corresponds generally to the zone E; also, the hypersensitive area II includes the enhancer element located at -1100 (13). The tissue specificity and hormonal regulation of the hypersensitive site I seem to support the putative regulatory role of the protein binding sites present within this region. In particular, the increase in hypersensitivity of Site E around -400 upon hormone withdrawal is consistent with the localization by 5'-deletion mapping of a site important for the repression of the gene between -426 and -295 (2). It is tempting to speculate that the footprint at -330 represents the binding of a "repressor". However, we have no direct evidence at present to support this hypothesis.

Multiple *trans*-acting factors are known to interact with the upstream regulatory elements of many eukaryotic promoters. It is believed generally that the complex regulation of an eukaryotic promoter is achieved through the functional interaction of these factors with each other and/or with the general transcription initiation factors such as TATA-box binding protein or, RNA polymerase II. Cooperative interactions between different upstream regulatory factors and the TATA-box binding protein recently have been reported (23,24). Different combinations of such interactions may result in specific patterns of expression of a gene. In the case of the ovalbumin gene, the presence of closely-spaced and closely-related binding sites in the upstream promoter area

raises the interesting possibility that a complex interplay of these proteins, the steroid receptor, and the core-promoter factors may establish the highly tissue-specific pattern of expression of this gene. To prove that such a scenario exists, one should functionally dissect the ovalbumin promoter by mutation and analyze its expression in homologous oviduct cells. A major obstacle is the lack of an appropriate cultured oviduct cell line which can be transfected with cloned DNA. It has been possible recently in some systems to reproduce some characteristics of the tissue specific activation of transcription in vitro (25,26). The procedures rely in large part on the careful purification of intact nuclei (25), and these preparations are difficult to assess with tissues such as oviducts where nucleases are particularly abundant. At present, we lack a cell-free hormone-regulatable transcription system to study directly the functional roles of the upstream promoter binding proteins *in vitro*. An alternative approach is to study the protein-DNA structure of the regulatory region of genes such as ovalbumin. To this end, we will attempt extensive purification of at least some of the tissue-specific regulatory components of the ovalbumin gene as a prelude to reconstitution of such a system in vitro.

ACKNOWLEDGEMENTS

We thank Serlina Robinson, Jessica Liu, and Craig Wilkinson for excellent technical assistance and Dr. Juan Codina-Salada for synthesizing the deoxyoligonucleotides. We also thank Carolyn Armijo for careful preparation of this manuscript. This work was supported by Public Health Service grant HD-08188 (BWO) and HD-17379 (MJT) from the National Institute of Health.

*To whom correspondence should be addressed

REFERENCES

1. Kaye, J.S., S. Pratt-Kaye, M. Bellard, G. Dretzen, F. Bellard, and P. Chambon (1986). *EMBO J.* 5: 277-285.
2. Gaub, M.-P., A. Dierich, D. Astinotti, I. Touitou, and P. Chambon (1987). *EMBO J.* 6: 2313-2320.
3. O'Malley, B.W., D.R. Roop, E.C. Lai, J.L. Nordstrom, J.F. Catterall, G.E. Swaneck, D.A. Colbert, M.-J. Tsai, A. Dugaiczky, and S.L.C. Woo (1979). *Recent Progress Hormone Res.* 35: 1-46.
4. Hager, L.J., G.S. McKnight and R.D. Palmiter (1980). *J. Biol. Chem.* 254: 7796-7800.
5. Tsai, S.Y., M.-J. Tsai, and B.W. O'Malley (1981). *Proc. Natl. Acad. Sci. USA* 78: 879-883.
6. Zarucki-Schulz, T., S.Y. Tsai, K. Itakura, X. Soberon, R.B. Wallace, M.-J. Tsai, S.L.C. Woo, and B.W. O'Malley (1982). *J. Biol. Chem.* 257: 11070-11077.

7. Knoll, B.J., T. Zarucki-Schultz, D.C. Dean, and B.W. O'Malley (1983). *Nucleic Acids Res.* 11: 6733-6754.
8. Pastorcic, M., H. Wang, A. Elbrecht, S.Y. Tsai, M.-J. Tsai, and B.W. O'Malley (1986). *Mol. Cell Biol.* 6:2784-2791.
9. Sagami, I., S.Y. Tsai, H. Wang, M.-J. Tsai, and B.W. O'Malley (1986). *Mol. Cell Biol.* 6: 4259-4267.
10. Tsai, S.Y., I. Sagami, H. Wang, M.-J. Tsai, and B.W. O'Malley (1987). *Cell* 50: 701-709.
11. Wang, L.H., S.Y. Tsai, I. Sagami, M.-J. Tsai, and B.W. O'Malley (1987). *J. Biol. Chem.* 262: 16080-16086.
12. Bagchi, M.K., S.Y. Tsai, M.-J. Tsai, and B.W. O'Malley (1987). *Mol. Cell Biol.* 7:4151-4158.
13. Bradshaw, M.S., M.-J. Tsai, and B.W. O'Malley (1988). *J. Biol. Chem.* 263: 8485-8490.
14. Dierich, A., M.-P. Gaub, J.-P. LePennec, D. Astinotti, and P. Chambon (1987). *EMBO J.* 6: 2305-2312.
15. Garner, M.M., and A. Revzin (1981). *Nucleic Acids Res.* 9: 3047-3060.
16. Briggs, M., J. Kadonaga, S. Bell, and R. Tjian (1986). *Science* 234: 47-52.
17. Hager, D., and K. Burgess (1980). *Anal. Biochem.* 109: 76-86.
18. Morrissey, J.H. (1981). *Anal. Biochem.* 117: 307-313.
19. Tora L., B. Gronemeyer, B. Turcotte, M.-P. Gaub, and P. Chambon (1988). *Nature* 333: 185-188.
20. Jantzen, H.-M., U. Strahle, B. Gloss, F. Steward, W. Schmid, M. Boshart, R. Mksiecek, and G. Schutz (1987). *Cell* 49: 29-38.
21. Slater, E.P., O. Rabenau, M. Karin, J.D. Baxter, and M. Beato (1985). *Mol. Cell Biol.* 5: 2984-2992.
22. Gerlinger, P., A. Krist, M. Lemeur, F. Perrin, M. Cochet, F. Gannon, D. Dupret, and P. Chambon (1982). *J. Mol. Biol.* 162: 345-364.
23. Sawadogo, M. and R.G. Roeder (1985). *Cell* 43: 165-175.
24. Horikoshi, M., M.F. Carey, H. Kakidani, and R.G. Roeder (1988). *Cell* 54: 665-669.
25. Gorski, K., M. Carmiero, and U. Schibler (1986). *Cell* 47: 767-776.
26. Corthesy, B., R. Hipkind, I. Theulaz, and W. Wahli (1988). *Science* 239: 1137-1139.