MOLECULAR AND CELLULAR BIOLOGY, Jan. 1990, p. 154–164
0270-7306/90/010154-11$02.00/0
Copyright © 1990, American Society for Microbiology
Vol. 10, No. 1

Developmental Regulation of Specific Protein Interactions with an Enhancerlike Binding Site Far Upstream from the Avian Very-Low-Density Apolipoprotein II Gene

PAMELA A. HOODLESS,[1,2] RABINDRA N. ROY,[1] AIMEE K. RYAN,[1,2] ROBERT J. G. HACHÉ,† MONIKA Z. VASA,[1,2] AND ROGER G. DEELEY[1,2,3]*

Department of Biochemistry,[1] Department of Pathology,[3] and Cancer Research Laboratories,[2] Queen's University, Kingston, Ontario K7L 3N6, Canada

Received 15 June 1989/Accepted 25 September 1989

Expression of the avian very-low-density apolipoprotein II (apoVLDLII) gene is completely dependent on estrogen and restricted to the liver. We have identified binding sites for nonhistone nuclear proteins located between −1.96 and −2.61 kilobases. One of these sites, located at −2.6 kilobases (designated site 1), was found to span an *MspI* site that becomes demethylated between days 7 and 9 of embryogenesis, the stage of development at which competence to express the apoVLDLII gene begins to be acquired. Levels of the factor(s) involved were high at day 7 of embryogenesis, decreased two- to threefold by days 9 to 11, and continued to decline more slowly until hatching. Furthermore, the mobility of the complex formed underwent a well-defined shift between days 11 to 13 of embryogenesis. Methylation interference studies showed that modification of the outer guanosines of the *MspI* site resulted in marked inhibition of the formation of the protein-DNA complex. Competition studies, fractionation of nuclear extracts, and tissue distribution indicated that the factor was not the avian homolog of hepatocyte nuclear factor 1, nuclear factor 1, or CCAAT/enhancer-binding protein (C/EBP). However, site 1 could compete for binding to an oligonucleotide, previously shown to be recognized by C/EBP, in a nonreciprocal fashion. These studies demonstrate that the sequence recognized by the protein includes a C/EBP consensus sequence but that elements in addition to the core enhancer motif are essential for binding.

Expression of genes that are regulated by steroid hormones, particularly the sex steroids, displays a tissue specificity that may be restricted to an individual cell type, even though the appropriate receptor may be present in additional tissues (reviewed in reference 3). In the case of the estrogen and progesterone receptors, unlike other members of the nuclear receptor superfamily, such as those for thyroid hormone, there is no indication of the existence of tissue-specific receptor types. This suggests that developmentally programmed differences in chromatin structure or the availability of *trans*-acting factors in addition to the receptor itself dictate the tissue specificity, and possibly efficiency, with which a particular hormonally responsive gene is expressed (36). In most cases, the extent to which tissue specificity of expression depends on the concerted effects of the receptor and other *trans*-acting factors or on the presence of *cis*-acting sequences in addition to the hormone response element has not been established. The estrogen-dependent activation of genes encoding major egg proteins provides a complex but experimentally accessible model with which to investigate the contribution that each of these mechanisms may make in determining tissue specificity.

In birds, transcription of the major yolk protein genes, for example, those specifying the vitellogenins and very-low-density apolipoprotein II (apoVLDLII), is strictly limited to the liver and completely dependent on estrogen (41). In the case of the apoVLDLII gene, the liver does not begin to display the ability to express the gene in response to treatment with estrogen until sometime between days 7 and 9 of embryogenesis, whereas the ability to activate expression of the vitellogenin genes is not acquired until 2 to 3 days later (11, 16, 18). Although dormant in roosters, chicks, and embryos, these genes can be activated in birds of either sex without a requirement for DNA replication (25). Despite this fact, the liver displays a propagatable, estrogen-priming effect that is reflected by a long-lasting alteration in the kinetics with which the genes respond to restimulation with the hormone (10, 15, 17, 26, 30).

Analyses of overall nuclease sensitivity and nuclear matrix association support the suggestion that the egg white and yolk protein genes are differentially segregated into either inactive or poised chromatin in the liver or oviduct, depending on the tissue in which they are activatable. However, analyses of the spectrum of nuclease-hypersensitive sites in and around both the vitellogenin and apoVLDLII genes indicates that the genes are marked in the tubular gland cells of the oviduct by a subset of the sites that are present in the liver (8, 20, 27). The accessibility of these sites suggests that the lack of activation of the genes in the oviduct may not simply be attributable to their sequestration in heterochromatin but may also be the consequence of either the absence of tissue-specific factors that are required for expression or the presence of negative regulators that prevent activation.

We have screened approximately 5 kilobases (kb) of the 5′-flanking region of the apoVLDLII gene for the presence of specific binding sites for nonhistone nuclear proteins. These studies have revealed protein interactions with a region extending from 2.0 to 2.6 kb upstream of the major transcriptional initiation site of the gene. Some of the factors that bind to this region are present in both liver and oviduct. However, one of the binding sites detected is recognized by a

* Corresponding author.
† Present address: Institüt für Molekularbiologie und Tumorforschung, Philipps Universität, 3550 Marburg, Federal Republic of Germany.

factor, or factors, present in liver that displays a restricted tissue distribution. The DNA sequence involved contains inverted copies of core enhancer sequences (40). It also includes an *Msp*I site that we have shown previously to become demethylated at the stage of development during which the liver begins to display competence to express the apoVLDLII gene (11). The studies presented also demonstrate that (i) the protein(s) binding to this site is not the avian homolog of previously characterized mammalian DNA-binding proteins that recognize related sequences, such as CCAAT/enhancer-binding protein (C/EBP) (24), nuclear factor 1 (NF-1) (6, 7, 29), and hepatocyte nuclear factor 1 (HNF-1) (14); (ii) although the protein involved is not C/EBP, the site competes in a nonreciprocal fashion with a sequence known to bind C/EBP; (iii) the expression of a protein(s) binding to the site changes markedly during embryogenesis; and (iv) the *Msp*I site itself is intimately involved in the interaction between protein and DNA.

MATERIALS AND METHODS

Isolation and labeling of DNA fragments. DNA from the 5′-flanking region of the ApoVLDLII gene was subcloned as an *Eco*RI fragment of 5.4 kb into pBR322. The fragment spanned from 800 nucleotides downstream of the major transcriptional start site to 4.6 kb upstream. DNA fragments described here were subcloned from this initial fragment. Oligonucleotides were synthesized by K. V. Deugau, using the phosphoramidite procedure and a Biosearch 3700 automated synthesizer, and were purified by thin-layer chromatography (1). Fragments and oligonucleotides used for binding and exonuclease III mapping were labeled at their 5′ ends with polynucleotide kinase and [γ-^{32}P]ATP. All fragments and oligonucleotides were made flush ended by fill-in synthesis with the Klenow fragment of *Escherichia coli* DNA polymerase before use.

Preparation of protein extracts. Crude nuclear extracts of nonhistone proteins were prepared from liver and oviduct by micrococcal nuclease digestion and extraction with 0.3 M NaCl or by extraction of intact nuclei with 0.4 M KCl essentially as described previously (6, 31). An additional 40% (wt/vol) ammonium sulfate precipitation was introduced to minimize nuclease contamination that was most apparent in liver preparations. In some cases, extracts were purified further by chromatography on reactive orange 14 agarose columns prepared as described previously (5). Proteins were eluted with step or linear gradients ranging from 0.1 to 1.0 M in NaCl.

Filter binding assays. Filter binding assays were carried out essentially as described by Riggs et al. (32). Various quantities of protein extracts (specified in the figure legends) were preincubated for 30 min at 4°C in 100 μl of buffer containing *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (HEPES; pH 8.3, 10 mM), NaCl (100 mM), $MgCl_2$ (10 mM), dithiothreitol (2 mM), bovine serum albumin (50 μg/ml), and *Alu*I-digested *E. coli* or pBR322 DNA (2 μg/ml). Labeled DNA fragments with flush ends (20 ng) were then added to the mix, and incubation continued for an additional 30 min. The mixture was then filtered through BA85 nitrocellulose filters (Schleicher & Schuell, Inc.). DNA fragments bound to protein that were selectively retained were eluted by vigorous shaking of the filter for 1 h at 37°C in digestion buffer containing Tris (pH 8.0, 50 mM), EDTA (10 mM), NaCl (50 mM), sodium dodecyl sulfate (0.25%), and proteinase K (200 μg/ml). The eluate was extracted with phenol-chloroform, and the DNA was precipitated with 2.5 volumes of ethanol. Fragments were analyzed on polyacrylamide gels (5%) run in Tris-borate buffer (pH 8.3) containing Tris (100 mM), borate (100 mM), and EDTA (2 mM).

Gel electrophoresis mobility shift assays. Incubation of end-labeled DNA fragments with nuclear extracts was performed as described for the filter binding assays, with incorporation of the following modifications. Incubations were carried out in 25 μl containing 0.5 ng of DNA fragment (specific activity, 9,000 to 12,000 cpm/ng). Assays using oligonucleotides were carried out in 25 μl containing 25 mM HEPES (pH 8.0), 12.5 mM $MgCl_2$, 100 mM KCl, 20% glycerol, and 2 mM dithiothreitol with 2.5 fmol of oligonucleotide (specific activity, 2×10^6 to 3×10^6 cpm/fmol) and a 1,500-fold excess of poly(dI-dC). Band shifting assays (19) were carried out by using polyacrylamide gels of various percentages that had been cast in Tris-acetate buffer (pH 7.2) containing Tris (20 mM), sodium acetate (10 mM), and EDTA (2 mM). Before use, the gels were preelectrophoresed for 1 h at 80 V. Electrophoresis was carried out at 80 V for periods ranging from 2 to 4 h. Gels were then dried and subjected to autoradiography.

Methylation interference experiments. Methylation interference was carried out as described by Sen and Baltimore (37). End-labeled DNA was partially methylated by dimethyl sulfate for 3 min at room temperature. Preparative binding reactions were carried out as described above, scaled up 10-fold. Free and bound DNA fragments were separated on a deionized 6% polyacrylamide gel, using the buffer conditions stated above. After electrophoresis, the gel was wrapped in plastic wrap and exposed to Kodak XAR-5 for 3 h at 4°C. The autoradiograph was used to locate bound and free fragments, which were then excised and electroeluted for 2 h to recover the DNA. Electroeluted DNA was incubated with proteinase K overnight (1 mg/ml) and then extensively extracted with phenol and chloroform, followed by ethanol precipitation. After cleavage with piperidine and lyophilization, the products were separated on a 15% polyacrylamide gel containing 8 M urea and subjected to autoradiography.

RESULTS

Screening for protein-binding sites in the 5′-flanking region of the apoVLDLII gene. Screening was carried out initially on a 5.4-kb *Eco*RI fragment extending 4.6 kb upstream of the major start site of the gene that also included the first exon and 800 nucleotides of the 5′-proximal intron. Binding studies were performed with liver extracts from normal day 17 embryos as well as from embryos that had been treated with estrogen for 48 h before sacrifice. The extracts were prepared either by salt extraction of intact nuclei with 0.4 M KCl or by micrococcal nuclease digestion to solubilize approximately 20% of the DNA, followed by extraction with 0.3 M NaCl. Various amounts of protein were incubated with end-labeled fragments from a *Hin*fI digest of plasmid 5.4/pBR322, which contained the 5.4-kb fragment inserted at the *Eco*RI site. This digest generated a total of 22 fragments, 12 of which were derived from the original cloned fragment. Because of the complexity of the mixture, filter retention assays were used as a preliminary method of screening. One fragment in the mixture was preferentially bound by all extracts. Restriction mapping and sequencing demonstrated that the fragment spanned the region between base pairs (bp) −2824 and −1959 relative to the major transcriptional initiation site of the gene. This region includes the 5′-flanking *Msp*I site that is demethylated during embryogenesis.

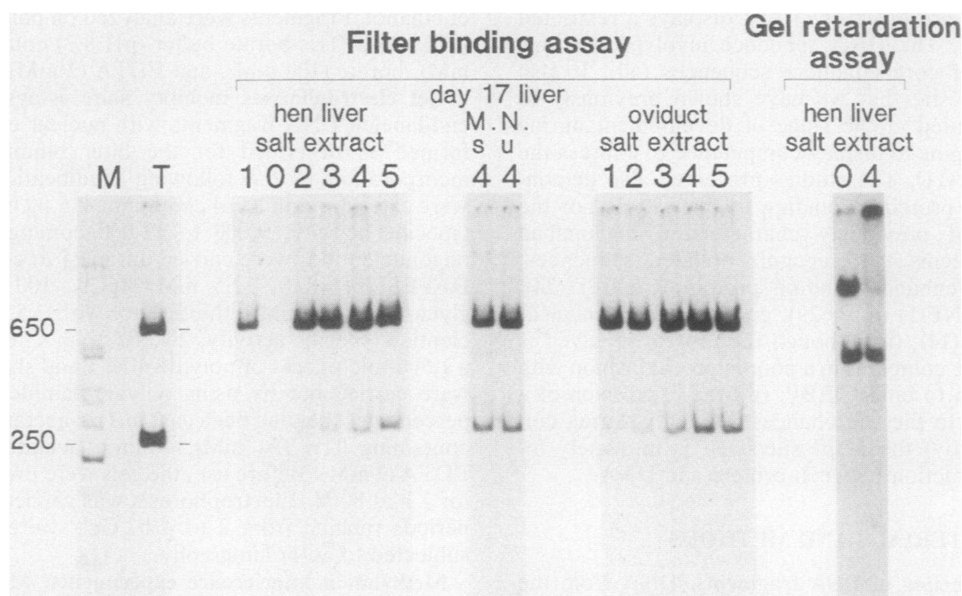

FIG. 1. Localization of protein binding to a region between 2.6 and 2.0 kb upstream of the transcriptional initiation site. The *Hin*fI fragment spanning the region between −2824 and −1959 bp was digested with *Msp*I, and the ends of the fragments were filled in, generating two fragments of approximately 650 and 250 bp (lane T). The ability of nuclear proteins to bind these fragments was assayed by filter binding or gel mobility shift assays. Nuclear proteins were either salt extracted from hen liver and oviduct nuclei or extracted after micrococcal nuclease digestion (M.N.) of embryonic liver nuclei from normal day 17 embryos (u) or from embryos treated with estrogen 48 h before sacrifice (s). Each assay contained 5 ng of labeled DNA fragments, 100 ng of *Alu*I-digested pBR322, and the amount of nuclear protein indicated (in micrograms) above each lane. Lane M, *Hin*fI-digested pBR322 markers.

To determine whether we could localize binding to one side or the other of the *Msp*I site, the fragment was digested with the enzyme and the ends of the two subfragments were filled in. They were then used for filter retention and gel retardation assays carried out with extracts from both liver and oviduct. Protein binding appeared to selectively take place on the larger of the two fragments (Fig. 1). No difference in binding ability could be detected between extracts from liver or oviduct nor was any dependence on the hormonal status of the bird apparent. However, under the initial conditions used for the gel retardation assay, the bound fragment was completely retained at the origin of electrophoresis, raising the possibility that it contained multiple protein-binding sites.

Exonuclease III mapping was carried out on a series of fragments spanning from a *Dde*I site at −2667 to the *Hin*fI site at −1959. Using this approach, we were able to tentatively identify boundaries of several binding sites. The mapping of one of these sites, designated site 1, indicated that it spanned the *Msp*I site at −2592. On the downstream side of the site, strong exonuclease III stops were detected 35 and 24 nucleotides from the *Msp*I recognition sequence. Mapping of the upstream boundary produced a more complex pattern of stops that showed a marked dependency on protein concentration, possibly with a consequence of protein stacking on the DNA. However, at low protein concentration, two strong stops were detected: one 5 to 6 nucleotides upstream of the *Msp*I site and the other 17 nucleotides downstream. Both of these boundaries occurred at a GG doublet within two copies of the sequence TGGT. The sequence of the site and the positions of the upstream and downstream boundaries are shown in Fig. 2A. A particularly notable feature of the site was the presence of inverted copies, spaced 10 nucleotides apart, of the sequence $^{TT}_{AG}TG$ $GT^{AA}_{TT}G$. Each of these elements had only one mismatch with the core enhancer sequence, $TGTGG^{TTT}_{AAA}G$ (40). In addition, elements within the site displayed similarities to other sequences recognized by previously characterized *trans*-acting factors (Fig. 2A).

Tissue specificity of binding to site 1. To examine the specificity of binding to site 1 in greater detail, an oligonucleotide corresponding to the sequence shown in Fig. 2A was synthesized and used in gel retardation assays, initially with nuclear extracts from both liver and oviduct (Fig. 2B). Although no differences in binding could be demonstrated when the complete 650-nucleotide fragment was used, there was a marked difference in the ability of extracts from the two tissues to bind to the oligonucleotide. The liver extracts formed readily detectable amounts of a single complex. In contrast, binding could be detected with the oviduct extracts only after prolonged exposure of the autoradiograph. This revealed very low levels of two distinct complexes, one of which had a mobility similar but not identical to that of the complex detected with liver extracts.

A clearer comparison of the complexes formed with extracts from the two tissues was obtained when both were fractionated by chromatography on reactive orange 14 columns. Two complexes (numbered 1 and 2 in Fig. 3) could be detected in fractions containing peak binding activity from both extracts. However, the liver extracts formed complex 1 almost exclusively, whereas oviduct extracts formed very small amounts of both complexes.

The oviduct was chosen for initial comparisons of binding to site 1 because of the coordinate manner with which both tissues respond to estrogen. These analyses were subsequently extended to blood, brain, and kidney. The levels of binding were markedly lower in both blood and brain than in liver, whereas kidney nuclear extracts displayed binding activities that approached hepatic levels. However, in all cases the mobility of the complex formed was indistinguish-

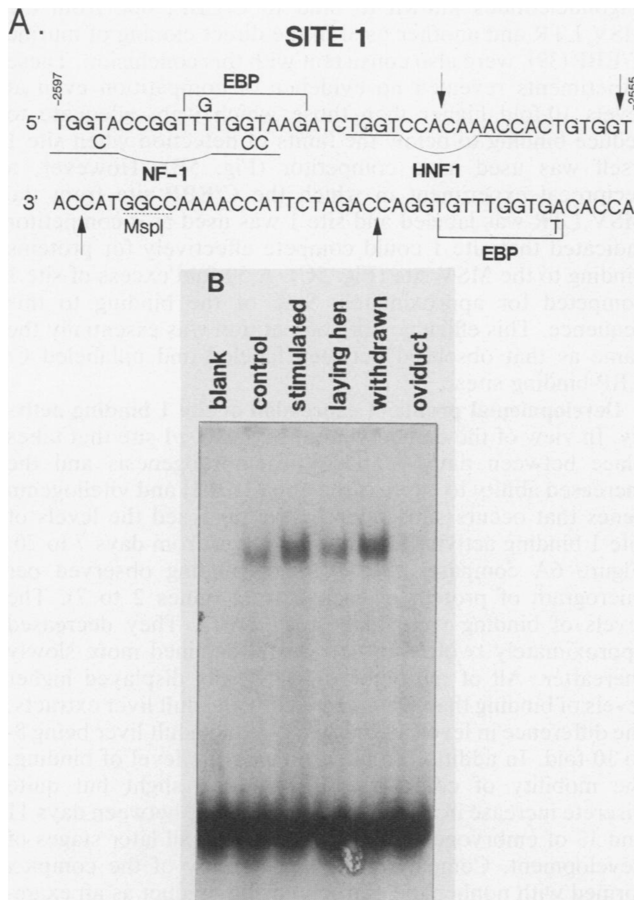

FIG. 2. (A) Sequence similarities between site 1 and the sequences recognized by previously described nuclear factors. Sequences of both strands of site 1 are shown, with the regions of similarity indicated as follows. Two inverted copies of a C/EBP core recognition sequence are indicated by square brackets. Mismatches are noted, with the consensus nucleotide located within the brackets. Two half-sites for an NF-1-binding site with a 7-nucleotide separation are underlined. Mismatches are again indicated, with the consensus nucleotide marked under the site. The diagram also shows an 8-nucleotide homology to an HNF-1-binding site (underlined) found upstream of the α_1-antitrypsin gene. Upstream (lower strand) and downstream (upper strand) boundaries determined by exonuclease III mapping (not shown) are indicated by arrows. (B) Binding of liver and oviduct nuclear extracts to synthetic site 1 DNA. End-labeled site 1 (10,000 cpm; 60 to 70 pg) was incubated without nuclear protein (blank) or with 1.0 μg of an extract from normal rooster liver (control), rooster liver treated with estrogen 48 h before sacrifice (stimulated), laying hen liver, rooster liver treated with estrogen 2 weeks before sacrifice (withdrawn), or oviduct. A 1,500-fold weight excess of poly(dI-dC) was added as a nonspecific competitor to each assay. Samples were then analyzed by gel mobility shift as described in Materials and Methods.

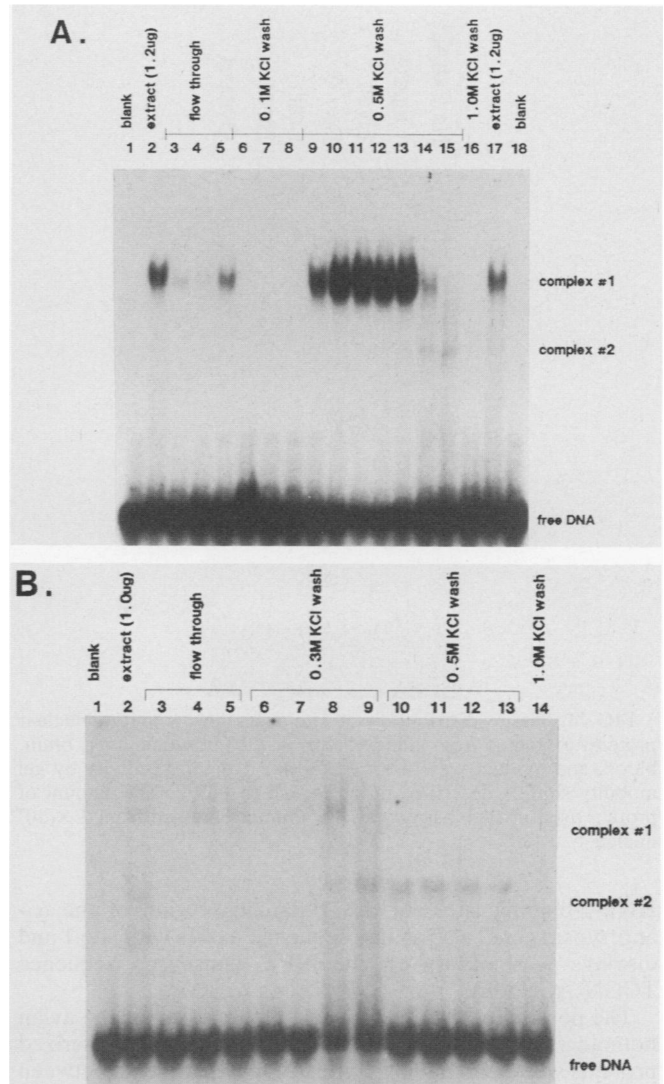

FIG. 3. Comparison of site 1 binding activity in liver and oviduct extracts fractionated by reactive orange 14 chromatography. (A) Nuclear extract from estrogen-treated rooster liver was fractionated on a reactive orange 14 column and assayed for binding to site 1, as described in the legend to Fig. 2. Assays contained no protein (lanes 1 and 18), 1.2 μg of crude nuclear extract (lanes 2 and 17), or 5 μl (<0.4 μg of protein) of column fractions from flowthrough (lanes 3 to 5), 0.1 M KCl wash (lanes 6 to 8), 0.5 M KCl wash (lanes 9 to 15), and 1.0 M KCl wash (lane 16). (B) Nuclear extract from oviduct was fractionated and assayed as described above. Assays contained no protein (lane 1), 1.0 μg of crude extract (lane 2) or 5 μl of column fractions (0.1 to 1.0 μg of protein) from flowthrough (lanes 3 to 5), 0.3 M KCl wash (lanes 6 to 9), 0.5 M KCl wash (lanes 10 to 13), or 1.0 M KCl wash (lane 14).

able from that observed with oviduct extracts and slightly slower than that obtained with adult liver (Fig. 4).

Comparison of site 1 binding specificity with specificities of previously identified mammalian and avian DNA-binding proteins. As pointed out above, the oligonucleotide corresponding to site 1 contains two motifs that are very similar to viral enhancer core sequences. A protein termed C/EBP has been detected in mammalian liver and very recently in several additional tissues (4). C/EBP has been shown to bind to the core sequences of several viral and cellular enhancers that contain variations of this core enhancer sequence (12, 24). Site 1 also contains motifs that are similar to sequences shown to interact with members of the NF-1 family. For example, the site contains repeats of the pentanucleotide TGGTA similar to the TG$^{G}_{C}$CA elements found in the NF-1-binding site flanking the chicken lysozyme gene (6). In addition, a liver-specific protein, HNF-1, has been shown to bind to sites in the promoters of several mammalian genes that display liver-specific expression (14, 22). More recently, evidence has been presented that it may also interact with the enhancers of at least some of these genes (35). One of the

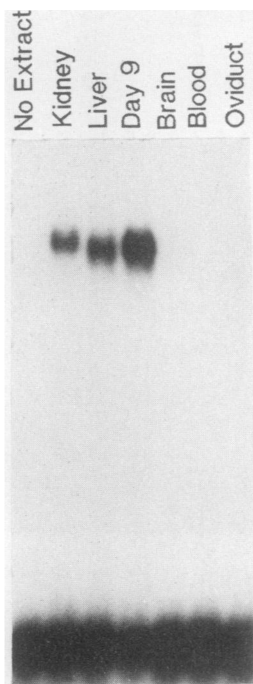

FIG. 4. Tissue distribution of site 1 binding activity. Nuclear proteins extracted from kidney, liver, day 9 embryonic liver, brain, blood, and oviduct were assayed for site 1 binding activity by gel mobility shift as described in the legend to Fig. 2. The amount of protein used in each assay was the amount obtained from 2×10^6 nuclei.

HNF-1-binding sites, in the 5'-flanking region of the α_1-antitrypsin gene, also shares sequence motifs with site 1 and displays a similarity to the NF-1 consensus sequence $TGGN_7ACC$ (38).

The possibility that site 1 was a binding site for the avian homolog(s) of one or more of these previously characterized proteins was examined by competition experiments between the site 1 oligonucleotide and sequences known to be binding sites for NF-1 (chicken lysozyme gene [6]), HNF-1 (beta fibrinogen [13]), and C/EBP (murine sarcoma virus [MSV] long terminal repeat [LTR] [24] and a sequence used by Vinson et al. for the direct cloning of murine C/EBP [39]). Examples of competition experiments are shown in Fig. 5.

Binding to the NF-1 site was analyzed in the presence of up to a 500-fold excess of cold site 1 oligonucleotide (Fig. 5A). No competition was detected by this or reciprocal experiments in which the site 1 oligonucleotide was labeled. Similar experiments with the HNF-1-binding site also revealed no relationship with the protein(s) binding to site 1 (data not shown). Results of subsequent methylation interference experiments described below were also inconsistent with the binding being attributable to HNF-1. However, the results obtained with the C/EBP-binding sites were more complex.

Since the C/EBP-binding site identified in the MSV LTR was so similar to two elements in site 1, we first compared the levels of binding detected with this oligonucleotide in both liver and oviduct in the absence of competition. In contrast to the results obtained with site 1, the levels of binding were severalfold higher in the oviduct than in the liver, suggesting that binding to site 1 was not attributable to C/EBP. Results of experiments in which site 1 was labeled and binding was carried out with an excess of each of two oligonucleotides known to bind to C/EBP, one from the MSV LTR and another used in the direct cloning of murine C/EBP (39), were also consistent with this conclusion. These experiments revealed no evidence of competition even at levels 10-fold higher than those which were adequate to reduce binding to below the limits of detection when site 1 itself was used as a competitor (Fig. 5B). However, a reciprocal experiment in which the C/EBP site from the MSV LTR was labeled and site 1 was used as a competitor indicated that site 1 could compete effectively for proteins binding to the MSV site (Fig. 5C). A 50-fold excess of site 1 competed for approximately 90% of the binding to this sequence. This efficiency of competition was essentially the same as that observed between labeled and unlabeled C/EBP-binding sites.

Developmental profile of expression of site 1 binding activity. In view of the demethylation of the *Msp*I site that takes place between days 7 and 9 of embryogenesis and the increased ability to express the apoVLDLII and vitellogenin genes that occurs subsequently, we analyzed the levels of site 1 binding activity in embryonic liver from days 7 to 20. Figure 6A compares the levels of binding observed per microgram of protein in each extract (lanes 2 to 7). The levels of binding were highest at day 7. They decreased approximately twofold by day 9 and declined more slowly thereafter. All of the embryonic extracts displayed higher levels of binding than were detected with adult liver extracts, the difference in levels between day 7 and adult liver being 8- to 10-fold. In addition to alterations in the level of binding, the mobility of complex 1 displayed a slight but quite discrete increase in mobility that took place between days 11 and 13 of embryogenesis and persisted at all later stages of development. Comparison of the mobility of the complex formed with nonhepatic extracts, using oviduct as an example, indicated that it was identical with that of the complex detected in the liver before day 13 (Fig. 6A).

Since site 1 displayed the potential to compete with the C/EBP-binding site of the MSV LTR, we carried out a second analysis to determine whether the proteins binding to both sites were coordinately regulated during development. The results indicated that protein(s) binding to the MSV LTR sequence displayed a similar but not identical profile of expression (Fig. 6B). As with binding to site 1, the levels of binding to the MSV LTR site decreased severalfold during embryogenesis. However, the profile was distinct from that for binding to site 1 in two respects: (i) levels of binding remain high from days 7 through 11, declining two- to threefold between days 11 and 13, and (ii) the mobility of the complex remained constant throughout embryogenesis. A similar analysis using an NF-1-binding site (data not shown) gave a constant level of binding throughout embryogenesis.

Identification of elements required for binding. The two downstream boundaries detected by exonuclease III mapping were located at 5' and 3' ends of the downstream C/EBP-like binding element. An upstream boundary was also detected between the two elements. To determine whether both upstream and downstream elements were required for the binding activity observed, the oligonucleotide was asymmetrically labeled and digested with *Bgl*II and the fragments were either filled in or trimmed with S1 nuclease. This generated two labeled upstream fragments, one of which lacked the 3'-terminal G of the upstream element after S1 treatment and the other of which included this residue plus 3 nucleotides on the 3' side of it. Both fragments were purified by electrophoresis and used for gel retardation assays (Fig. 7). Essentially no diminution in

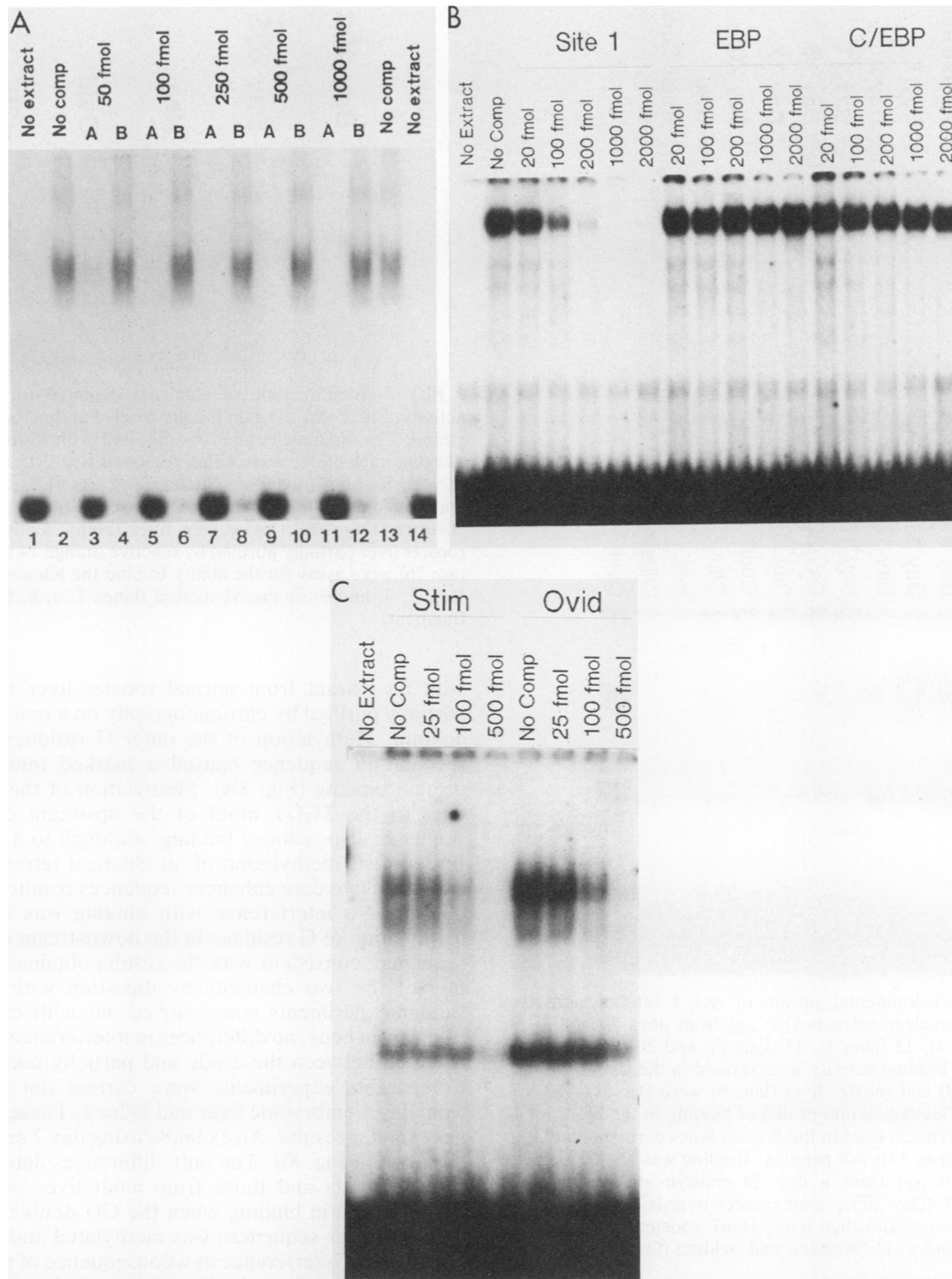

FIG. 5. Competition binding of site 1 with an NF-1-binding site and site 1 with two C/EBP-binding sites. (A) Partially purified liver nuclear extract was incubated with a labeled NF-1-binding site in the absence (lanes 2 and 13) or presence of increasing amounts (indicated) of unlabeled NF-1-binding site (lanes 3, 5, 7, 9, and 11) or unlabeled site 1 (lanes 4, 6, 8, 10, and 12). (B) Crude liver extract from a 48-h estrogen-treated rooster was incubated with labeled site 1 in the absence (No Comp) or presence of 20, 100, 200, 1,000, and 2,000 fmol of unlabeled site 1, EBP (C/EBP site from the MSV LTR [23]), and C/EBP (39). (C) Crude estrogen-treated liver extract (Stim) and oviduct extract (Ovid) were incubated with labeled C/EBP-binding site from the MSV LTR in the absence (No Comp) or presence of 25, 100, and 500 fmol of unlabeled site 1.

binding to the fragment that had been filled in was detected. Furthermore, the complex formed with this fragment displayed the same differences in mobility when extracts from liver and oviduct were compared. In contrast, binding to the fragment that had been treated with S1 was severely reduced.

The upstream half of the binding site was also used to examine the possibility that the inability of the C/EBP oligonucleotide, which contained a single core sequence, to compete with intact site 1 was attributable to cooperativity between the two C/EBP-like elements present. The results of competition experiments with the upstream half of the site

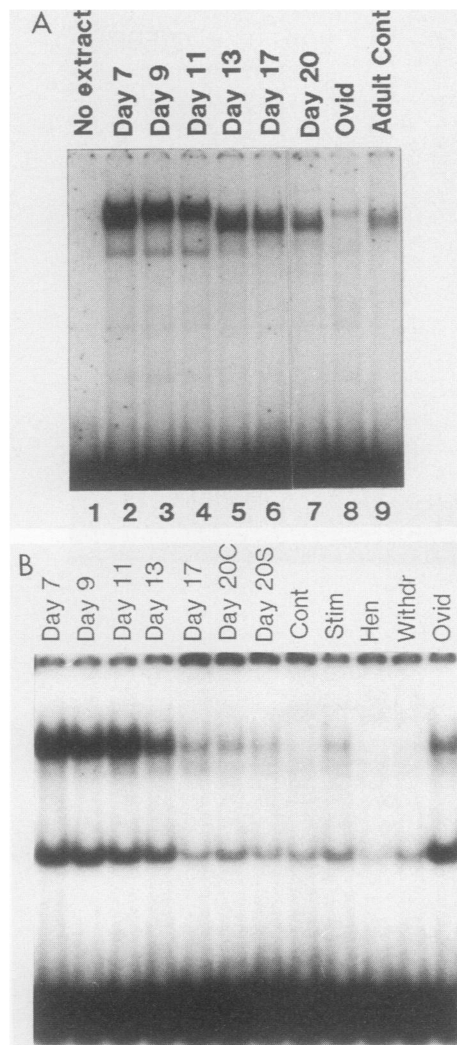

FIG. 6. (A) Developmental profile of site 1 binding activity. Embryonic liver nuclear extracts (1.0 μg) from days 7 (lane 2), 9 (lane 3), 11 (lane 4), 13 (lane 5), 17 (lane 6), and 20 (lane 7) were assayed for site 1 binding activity as described in the legend to Fig. 2. Oviduct (lane 8) and rooster liver (lane 9) were also assayed for comparison. (B) Developmental profile of binding to the MSV LTR C/EBP site. The extracts used in the first six lanes correspond to the extracts used in lanes 2 to 7 in panel A. Binding was also examined with extracts (1.0 μg) from a day 20 embryo stimulated with estrogen at day 18 (Day 20S), adult rooster liver (Cont), estrogen-treated rooster liver (Stim), hen liver (Hen), rooster liver 2 weeks after hormone withdrawal (Withdr), and oviduct (Ovid).

did not differ significantly from those obtained previously with the whole site (data not shown). This finding supports the conclusion that cooperativity between the C/EBP core elements cannot explain the nonreciprocal competition data and suggests that sequences, in addition to a single copy of this element, are required to define tissue-specific binding to site 1.

In view of the differences in mobility between complexes formed with adult liver extracts and extracts from early embryonic liver and nonhepatic tissues, we used methylation interference experiments to probe for alterations in DNA-protein contacts that would be indicative of a change in the conformation or composition of the protein moiety of the complex. Interference experiments were carried out first

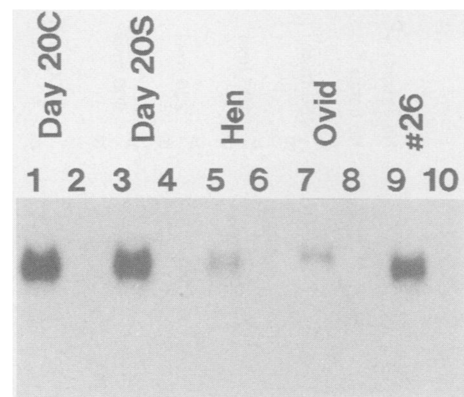

FIG. 7. Identification of elements essential for site 1 binding activity. Site 1 was asymmetrically labeled at the 5′ end of the upper strand. The oligonucleotide was digested with *Bgl*II, and the overhanging nucleotides were either removed with S1 nuclease or filled in with Klenow fragment. Nuclear extracts (1.0 μg) from normal embryonic day 20 liver (Day 20C), estrogen-treated embryonic day 20 liver (Day 20S), laying hen liver (Hen), oviduct (Ovid), and rooster liver partially purified by reactive orange 14 chromatography (no. 26) were assay for the ability to bind the Klenow-treated (lanes 1, 3, 5, 7, and 9) or the S1-treated (lanes 2, 4, 6, 8, and 10) *Bgl*II fragment.

with an extract from normal rooster liver that had been partially purified by chromatography on a reactive orange 14 column. Methylation of the outer G residues of the *Msp*I recognition sequence caused a marked interference with protein binding (Fig. 8A). Methylation of the guanine residues in the TGGT motif of the upstream core enhancer sequence also reduced binding, although to a lesser extent. In contrast, methylation of an identical tetranucleotide between the two core enhancer sequences resulted in increased binding. No interference with binding was detected after methylation of G residues in the downstream enhancer core sequence, consistent with the results obtained after separation of the two elements by digestion with *Bgl*II. When similar experiments were carried out with crude liver extracts from hens, no differences in interference patterns were detected between the crude and partially purified extracts. Comparable experiments were carried out with extracts from day 7 embryonic liver and kidney. These extracts gave very similar results. An example using day 7 embryonic liver is shown in Fig. 8B. The only differences detected between these extracts and those from adult liver were a lack of enhancement in binding when the GG doublet between the core enhancer sequences was methylated and a decrease in the extent of interference as a consequence of methylation of the upper-strand outer G residue of the *Msp*I site. No contact points were detected in the downstream core sequence with any of the extracts tested.

DISCUSSION

The ability to activate the apoVLDLII gene in response to estrogen provides the earliest detectable indication that the liver has acquired the competence to express at least one of several major yolk protein genes that would normally remain dormant until the onset of vitellogenesis in the hen. The competence to express the apoVLDLII gene and other members of this group, such as the vitellogenin genes, as well as the extent to which they respond, has been correlated with increases in the hormone-dependent inducibility of the

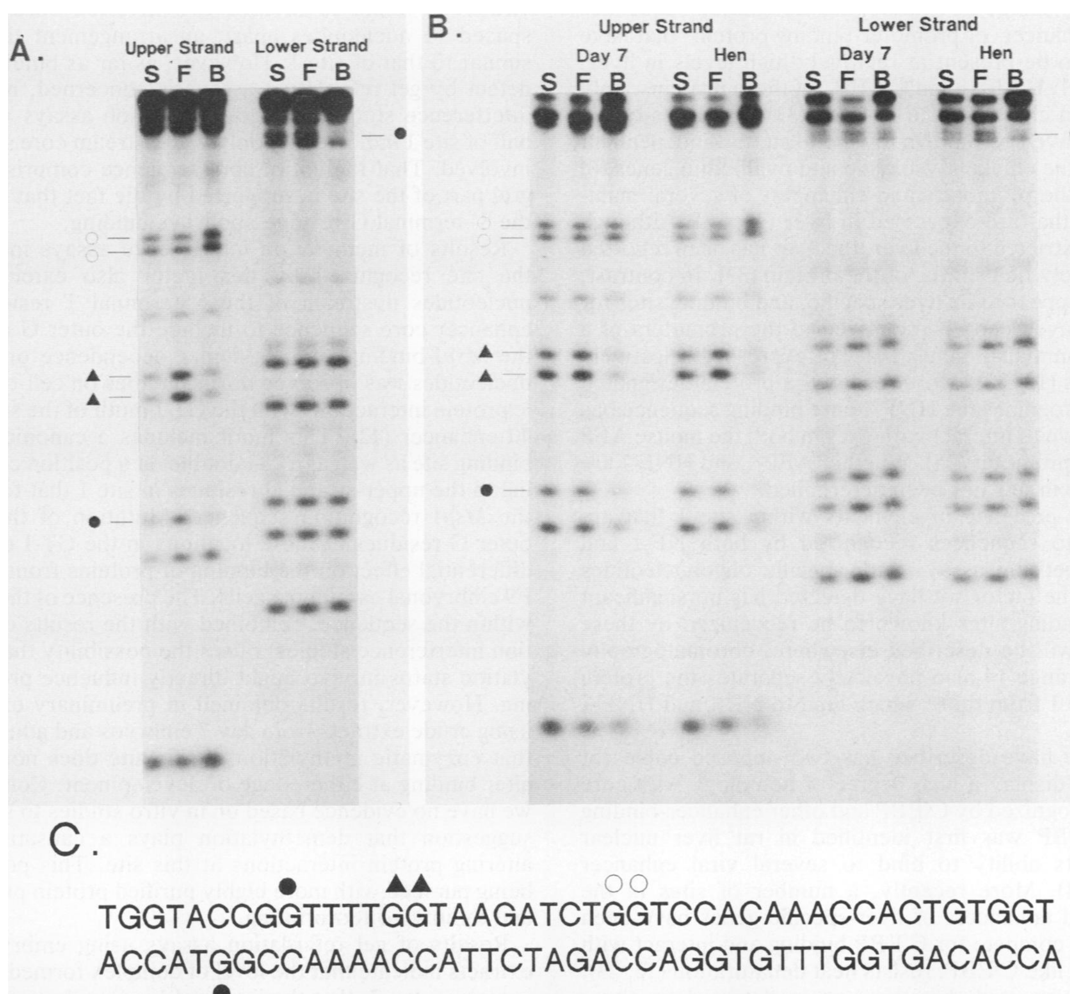

FIG. 8. Methylation interference of protein binding to site 1. (A) Analysis with partially purified rooster liver extract of the ability of methylation to interfere with protein binding to site 1. Partially methylated site 1 labeled on either the upper or lower strand (lanes S) was incubated with protein. Free (lanes F) and bound (lanes B) DNA was separated by gel mobility shift. Symbols: ●, methylated guanosines showing marked interference with protein binding; ▲, reduced binding; ○, enhanced binding. (B) Methylation interference assay with laying hen liver and embryonic day 7 liver nuclear extracts, performed as described above. (C) Sequence of site 1 indicating points showing methylation interference and enhancement of protein binding as indicated above.

hepatic estrogen receptor (16, 17). Induction of the receptor is first detectable at approximately days 9 to 10 of embryogenesis (16, 17, 21). The extent to which it can be induced then increases markedly during later stages of development. Thus, the developmental stage at which individual genes can be activated, and their levels of expression, could simply be a function of the competence of the liver to express estrogen receptor and a reflection of its relative affinities for individual estrogen response elements. However, previous analyses of DNase I-hypersensitive sites (20, 27) and methylation patterns (11) upstream of the 5′ end of the apoVLDLII gene have revealed developmental alterations in chromatin structure that take place after day 7 of embryogenesis and that extend at least 3.0 kb upstream of the 5′ end of the apoVLDLII gene. These changes occur in the absence of activation of the gene by estrogen. They raise the possibility that developmentally programmed alterations in expression of factors required for efficient transcription of the gene, in addition to the receptor, could also be taking place during this period.

The fragment to which binding was detected with crude nuclear extracts from both liver and oviduct begins approximately 100 to 150 nucleotides upstream of a liver-specific DNase I-hypersensitive site at approximately −1.8 kb. This site is very prominent in rooster liver but is not detectable in day 7 embryos (20). The region extends upstream to the *Msp*I site at −2592. It is probable that this fragment was detected in the initial screen with extracts from both liver and oviduct because of the presence within it of multiple protein-binding sites. In addition to specific protein binding to the region around the *Msp*I site, at least two other sites, which will be described elsewhere, were detected by exonuclease III mapping. Factors binding to these sites are present in both liver and oviduct and account for the lack of tissue specificity observed when the intact 650-nucleotide fragment was used.

We were particularly interested in protein binding to the region around the *Msp*I site because of the temporal correlation between demethylation of this site and acquisition of the ability to activate the apoVLDLII gene. The boundaries detected by exonuclease III mapping of site 1, together with particular features of its sequence, suggested that it could

contain closely linked or overlapping recognition sequences for several enhancer- or promoter-binding proteins that have been shown to be present at relatively high levels in liver, including NF-1, HNF-1, and C/EBP. Of these proteins, only NF-1 has been characterized in birds (33). Members of the NF-1 family have been shown to bind to steroid-independent enhancers of the chicken lysozyme and ovalbumin genes (6) as well as to the promoters and enhancers of several mammalian genes that are expressed in liver (23, 28). Although NF-1 is not restricted to the liver, the liver has been reported to be a relatively rich source of the protein (33). In contrast, HNF-1 does appear to be liver specific, and binding sites for this protein have been characterized in the promoters of a number of mammalian genes that are expressed selectively in hepatocytes (13, 22). More recently, a protein designated AFP-1 that recognizes the HNF-1 core binding sequence has also been shown to interact with sites in both the mouse AFP enhancer and promoter (35). Whether AFP-1 and HNF-1 are the same protein has not been determined.

Despite the presence of elements within site 1 that are very similar to sequences recognized by both NF-1 and HNF-1, competition assays with specific oligonucleotides indicate that the factor we have detected has no significant affinity for binding sites known to be recognized by these proteins. As will be described elsewhere, chromatography on reactive orange 14 also physically separates the protein binding to site 1 from those which bind to NF-1 and HNF-1 sites.

The site we have described has two opposed copies of elements that display a high degree of homology with core sequences recognized by C/EBP and other enhancer-binding proteins. C/EBP was first identified in rat liver nuclear extracts by its ability to bind to several viral enhancer sequences (24). More recently, a number of sites in the promoters and enhancers of liver-specific genes have been identified that compete for C/EBP binding and interact with a protein that, like C/EBP, resists heat denaturation (12, 23). It has been suggested that the protein detected by these analyses is either C/EBP or a member of a C/EBP-like family.

Several lines of evidence presented here indicate that the binding activity we describe is not attributable to C/EBP: (i) fractions from a reactive orange 14 column that contain the peak of site 1 binding activity display little or no binding to a C/EBP-binding site from MSV, (ii) in contrast to the factor binding to site 1, the factor binding to the C/EBP oligonucleotide is present in the oviduct at concentrations severalfold higher than in the liver, and (iii) the developmental regulation of the factor(s) binding to site 1 is distinct from that of proteins binding to a known C/EBP-binding site. In experiments not described here, we have also determined that the thermostability of the factor(s) binding to site 1 differs from that reported for C/EBP.

Although the evidence presented indicates that the factor we have identified is not the avian homolog of C/EBP, the site to which it binds can clearly compete with a known C/EBP-binding site, despite the fact that the reverse is not the case. In a very recent study, a heat-resistant avian factor was identified in crude liver nuclear extracts that binds to a number of mammalian C/EBP sites as well as closely related sites in the LTRs of avian retroviruses (34). The protein involved has been identified as the presumptive avian homolog of rodent C/EBP. The consensus core binding sequence that has been proposed for this protein is a perfect match for the upstream element in site 1. Interestingly, binding of this avian protein to the LTR of Rous sarcoma virus was found to involve two inverted core sequences spaced 11 nucleotides apart, an arrangement that is very similar to that of site 1. However, as far as binding that we detect by gel retardation assays is concerned, methylation interference studies and gel retardation assays with the 5′ half of site 1 indicate that only the upstream core sequence is involved. That the entire core sequence comprises an integral part of the site is supported by the fact that removal of the 3′-terminal G residue abolishes binding.

Results of methylation interference assays indicate that the site recognized by this factor also extends several nucleotides upstream of the 5′-terminal T residue of the enhancer core sequence to include the outer G residues of the *Msp*I-binding site. A similar dependence on upstream nucleotides was observed during studies on cell-type-specific protein interactions with the GT-I motif of the simian virus 40 enhancer (42). This motif includes a canonical C/EBP-binding site as well as a GG doublet at a position corresponding to the upper-strand G residues in site 1 that form part of the *Msp*I recognition sequence. Mutation of the inner or outer G residues at these locations in the GT-I motif has a differential effect on the binding of proteins from HeLa and F9 embryonal carcinoma cells. The presence of the *Msp*I site within the sequence, combined with the results of methylation interference studies, raises the possibility that its methylation status in vivo might directly influence protein binding. However, results obtained in preliminary experiments using crude extracts from day 7 embryos and adults indicate that enzymatic methylation of this site does not markedly alter binding at either stage of development. Consequently, we have no evidence based on in vitro studies to support the suggestion that demethylation plays a causative role in altering protein interactions at this site. This possibility is being pursued with more highly purified protein preparations and by in vivo footprinting.

Results of gel retardation assays using embryonic liver extracts indicate that the level of complex formed with site 1 is high at day 7, that the levels of binding decrease severalfold over the next 3 to 4 days, and that the mobility of the complex increases between days 11 to 13. The ability to activate the apoVLDLII gene is acquired at the beginning of this period, and the efficiency with which the gene can be expressed increases markedly between days 10 and 13 (16, 18). The mobility of the complex before day 13 is similar to the mobility of the complexes detected in extracts from nonhepatic tissues. Results of methylation interference assays indicate no qualitative differences in the binding between the lower- and higher-mobility complexes formed with embryonic and adult liver extracts, respectively. The shift in mobility of complexes formed with site 1 during liver development is similar in some respects to recent data reported during analyses of proteins interacting with liver-specific promoters in differentiated and dedifferentiated rat hepatoma cell lines (2, 9). These studies revealed a switch between factors binding to identical sites upon loss or acquisition of the differentiated phenotype that was initially detected as a slight shift in mobility in gel retardation assays. Despite the fact that different proteins appear to be involved, methylation interference assays revealed only slight or no alterations in binding properties.

The decrease in the level of binding to site 1 between days 7 and 11 is inconsistent with the factors identified at this time as playing a positive role in determining expression of the gene. However, it is possible that they could play a role either in establishment of the committed state or in suppression of expression of certain liver-specific functions. The

decline in levels of these factors in the liver and the shift in mobility of the complexes formed at later stages might then reflect a switch in binding to positive *trans*-acting factors. Our analyses using the C/EBP-binding site from the MSV LTR suggested that C/EBP might display a developmental profile in the liver similar to that observed with the factor binding to site 1, thus making it unlikely that C/EBP was the hypothetical positive factor. However, since completion of these studies, the developmental regulation of rodent C/EBP has been described with use of both cDNA and antibody probes (4). This study indicates that C/EBP expression increases late during embryogenisis and thus displays a developmental profile that could be consistent with such a role. These data also imply, unless the developmental profile of expression differs unexpectedly between birds and rodents, that the factor we have detected by gel retardation assays with the MSV C/EBP-binding site is distinct not only from the factor that binds to site 1 but also from C/EBP.

ACKNOWLEDGMENTS

We are grateful to Marilyn McCallum for excellent secretarial assistance and Ruth Burtch for technical assistance.

This research was supported by grants from the Medical Research Council of Canada. R.G. Deeley is the recipient of a Medical Research Council development grant. P.A. Hoodless and A.K. Ryan are the recipients of Medical Research Council studentships.

LITERATURE CITED

1. **Archer, T. K., S.-P. Tam, K. V. Deugau, and R. G. Deeley.** 1985. Apolipoprotein C-II mRNA levels in primate liver. J. Biol. Chem. **260**:1676–1681.
2. **Baumhueter, S., G. Courtois, and G. R. Crabtree.** 1988. A variant nuclear protein in dedifferentiated hepatoma cells binds to the same functional sequences in the β fibrinogen gene promoter as HNF-1. EMBO J. **7**:2485–2493.
3. **Beato, M.** 1989. Gene regulation by steroid hormones. Cell **56**:335–344.
4. **Birkenmeier, E. H., B. Gwynn, S. Howard, J. Jerry, J. I. Gordon, W. H. Landschulz, and S. L. McKnight.** 1989. Tissue-specific expression, developmental regulation, and genetic mapping of the gene encoding CCAAT/enhancer binding protein. Genes Dev. **3**:1146–1156.
5. **Bond, J. P., and A. C. Notides.** 1987. Estrogen receptor purification by affinity chromatography using an orange triazine dye. Anal. Biochem. **163**:385–390.
6. **Borgmeyer, U., J. Nowock, and A. E. Sippel.** 1984. The TGGCA-binding protein: a eukaryotic nuclear protein recognizing a symmetrical sequence on double-stranded linear DNA. Nucleic Acids Res. **12**:4295–4311.
7. **Bradshaw, M. S., M.-J. Tsai, and B. W. O'Malley.** 1988. A far upstream ovalbumin enhancer binds nuclear factor-1-like factor. J. Biol. Chem. **263**:8485–8490.
8. **Burch, J. B. E., and H. Weintraub.** 1983. Temporal order of chromatin structural changes associated with activation of the major chicken vitellogenin gene. Cell **33**:65–76.
9. **Cereghini, S., M. Blumenfeld, and M. Yaniv.** 1988. A liver-specific factor essential for albumin transcription differs between differentiated and dedifferentiated rat hepatoma cells. Genes Dev. **2**:957–974.
10. **Codina-Salada, J., J. P. Moore, and L. Chan.** 1983. Kinetics of primary and secondary stimulation of the mRNA for apoVLDL-II, a major yolk protein, in the cockerel liver by estrogen. Endocrinology **113**:1158–1163.
11. **Colgan, V., A. Elbrecht, P. Goldman, C. B. Lazier, and R. G. Deeley.** 1982. The avian apoprotein II very low density lipoprotein gene: methylation patterns of 5′ and 3′ flanking regions during development and following induction by estrogen. J. Biol. Chem. **257**:14453–14460.
12. **Costa, R. H., D. R. Grayson, K. G. Xanthopoulos, and J. E. Darnell, Jr.** 1988. A liver-specific DNA-binding protein recognizes multiple nucleotide sites in regulatory regions of transthyretin, α_1-antitrypsin, albumin, and simian virus 40 genes. Proc. Natl. Acad. Sci. USA **85**:3840–3844.
13. **Courtois, G., S. Baumhueter, and G. R. Crabtree.** 1988. Purified hepatocyte nuclear factor 1 interacts with a family of hepatocyte-specific promoters. Proc. Natl. Acad. Sci. USA **85**:7937–7941.
14. **Courtois, G., J. G. Morgan, L. A. Campbel, G. Gourel, and G. R. Crabtree.** 1987. Interaction of a liver-specific nuclear factor with the fibrinogen and α_1-antitrypsin promoters. Science **238**:688–692.
15. **Deeley, R. G., D. S. Udell, A. T. H. Burns, J. I. Gordon, and R. F. Goldberger.** 1977. Kinetics of avian vitellogenin messenger RNA induction. J. Biol. Chem. **252**:7913–7915.
16. **Elbrecht, A., C. B. Lazier, A. A. Protter, and D. L. Williams.** 1984. Independent developmental programs for two estrogen-regulated genes. Science **225**:639–641.
17. **Evans, M. I., P. J. O'Malley, A. Krust, and J. B. E. Burch.** 1987. Developmental regulation of the estrogen receptor and the estrogen responsiveness of five yolk protein genes in the avian liver. Proc. Natl. Acad. Sci. USA **84**:8493–8497.
18. **Evans, M. I., R. Silva, and J. B. E. Burch.** 1988. Isolation of chicken vitellogenin I and III cDNAs and the developmental regulation of five estrogen-responsive genes in the embryonic liver. Genes Dev. **2**:116–124.
19. **Fried, M., and D. M. Crothers.** 1981. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. **9**:6505–6525.
20. **Haché, R. J. G., and R. G. Deeley.** 1988. Organization, sequence and nuclease hypersensitivity of repetitive elements flanking the chicken apoVLDLII gene: extended sequence similarity to elements flanking the chicken vitellogenin gene. Nucleic Acids Res. **16**:97–113.
21. **Haché, R. J. G., S.-P. Tam, A. Cochrane, M. Nesheim, and R. G. Deeley.** 1987. Long-term effects of estrogen on avian liver: estrogen-inducible switch in expression of nuclear, hormone-binding proteins. Mol. Cell. Biol. **7**:3538–3547.
22. **Hardon, E. M., M. Frain, G. Paonessa, and R. Cortese.** 1988. Two distinct factors interact with the promoter regions of several liver-specific genes. EMBO J. **7**:1711–1719.
23. **Herbst, R. S., N. Friedman, J. E. Darnell, Jr., and L. E. Babiss.** 1989. Positive and negative regulatory elements in the mouse albumin enhancer. Proc. Natl. Acad. Sci. USA **86**:1553–1557.
24. **Johnson, P. F., W. H. Landschulz, B. J. Graves, and S. L. McKnight.** 1987. Identification of a rat liver nuclear protein that binds to the enhancer core element of three animal viruses. Genes Dev. **1**:133–146.
25. **Jost, J.-P., R. Keller, and C. Dierks-Ventling.** 1973. Deoxyribonucleic acid and ribonucleic acid synthesis during phosvitin induction by 17β-estradiol in immature chicks. J. Biol. Chem. **248**:5262–5266.
26. **Jost, J.-P., T. Ohno, S. Panyim, and A. R. Schuerch.** 1978. Appearance of vitellogenin mRNA sequences and rate of vitellogenin synthesis in chicken liver following primary and secondary stimulation by 17β-estradiol. Eur. J. Biochem. **84**:355–361.
27. **Kok, K., L. Snippe, G. Ab, and M. Gruber.** 1985. Nuclease-hypersensitive sites in chromatin of the estrogen-inducible apoVLDL II gene of chicken. Nucleic Acids Res. **13**:5189–5202.
28. **Lichtsteiner, S., J. Wuarin, and U. Schibler.** 1987. The interplay of DNA-binding proteins on the promoter of the mouse albumin gene. Cell **51**:963–973.
29. **Nagata, K., R. A. Guggenheimer, T. Enomoto, J. H. Lichy, and J. Hurwitz.** 1982. Adenovirus DNA replication *in vitro*: identification of a host factor that stimulates synthesis of the preterminal protein-dCMP complex. Proc. Natl. Acad. Sci. USA **79**:6438–6442.
30. **Noteborn, M. H. M., O. Bakker, M. A. W. de Jonge, M. Gruber, and G. Ab.** 1986. Differential estrogen responsiveness of the vitellogenin and apo very low density lipoprotein II genes in the rooster liver. J. Steroid Biochem. **24**:281–285.
31. **Nowock, J., and A. E. Sippel.** 1982. Specific protein-DNA interaction at four sites flanking the chicken lysozyme gene. Cell **30**:607–615.

32. **Riggs, A. D., H. Suzuki, and S. Bourgeois.** 1970. *lac* repressor-operator interaction. J. Mol. Biol. **48**:67–83.
33. **Rupp, R. A. W., and A. E. Sippel.** 1987. Chicken liver TGGCA protein purified by preparative mobility shift electrophoresis (PMSE) shows a 36.8 to 29.8 kd microheterogeneity. Nucleic Acids Res. **515**:9707–9726.
34. **Ryden, T. A., and K. Beemon.** 1989. Avian retroviral long terminal repeats bind CCAAT/enhancer-binding protein. Mol. Cell. Biol. **9**:1155–1164.
35. **Sawadaishi, K., T. Morinaga, and T. Tamaoki.** 1988. Interaction of a hepatoma-specific nuclear factor with transcription-regulatory sequences of the human α-fetoprotein and albumin genes. Mol. Cell. Biol. **8**:5179–5187.
36. **Schüle, R., M. Muller, C. Kaltschmidt, and R. Renkawitz.** 1988. Many transcription factors interact synergistically with steroid receptors. Science **242**:1418–1420.
37. **Sen, R., and D. Baltimore.** 1986. Multiple nuclear factors interact with the immunoglobulin enhancer sequences. Cell **46**:705–716.
38. **Shen, R. F., R. N. Sifers, H. Wang, C. Hardick, S. Y. Tsai, and S. L. C. Woo.** 1987. Tissue-specific expression of the human α_1-antitrypsin gene is controlled by multiple cis-regulatory elements. Nucleic Acids Res. **15**:8399–8415.
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. **Weiher, H., M. König, and P. Gruss.** 1983. Multiple point mutations affecting the simian virus 40 enhancer. Science **219**:626–631.
41. **Wiskocil, R., P. Bensky, W. Dower, R. F. Goldberger, J. I. Gordon, and R. G. Deeley.** 1980. Coordinate regulation of two estrogen-dependent genes in avian liver. Proc. Natl. Acad. Sci. USA **77**:4474–4478.
42. **Xiao, J.-H., I. Davidson, M. Macchi, R. Rosales, M. Vigneron, A. Staub, and P. Chambon.** 1987. In vitro binding of several cell-specific and ubiquitous nuclear proteins to the GT-I motif of the SV40 enhancer. Genes Dev. **1**:794–807.