MOLECULAR AND CELLULAR BIOLOGY, May 1988, p. 2070–2077
0270-7306/88/052070-08$02.00/0

Vol. 8, No. 5

The COUP Transcription Factor Binds to an Upstream Promoter Element of the Rat Insulin II Gene

YOUNG-PING HWUNG, DAVID T. CROWE, LEE-HO WANG, SOPHIA Y. TSAI, AND MING-JER TSAI*

Department of Cell Biology, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030

Received 1 December 1987/Accepted 17 February 1988

Band-shifting and DNase I-footprinting assays have been used to study the *trans*-acting factor(s) binding to an important promoter element (−53 to −46 relative to the transcription start) of the rat insulin II gene. A binding activity which footprints a region between −60 and −40 was found in both HIT, a hamster insulinoma cell line, and HeLa cells. A mutation within this region which drastically decreases promoter activity in vivo also greatly reduces binding activity in vitro. This binding activity was purified from HeLa cells and identified by competition and renaturation analyses as being the same as the COUP (chicken ovalbumin upstream promoter) transcription factor, a DNA-binding protein required for efficient transcription of the ovalbumin gene in vitro. Interestingly, the binding sequences of the COUP transcription factor in the ovalbumin and the insulin promoters have only limited similarities.

Accumulating evidence suggests that the regulation of gene expression is mediated by the interaction of *cis* elements and *trans*-acting factors (1, 6, 10, 14, 19, 21, 23, 26). Many sequence elements, including the TATA box, upstream promoter elements, enhancers, and silencers, have been identified as important for the expression of a large number of eucaryotic genes transcribed by RNA polymerase II. Similarly, a number of *trans*-acting factors which bind to these regulatory elements have been identified. Nevertheless, a critical question remains as to the molecular mechanism by which these *trans*-acting factors regulate transcription. A clear understanding requires the purification and characterization of these factors.

The insulin gene is expressed only in pancreatic β cells, thus providing a good model system for the study of tissue-specific gene expression. In rats, there are two nonallelic insulin genes which are highly homologous in flanking as well as coding regions (17) and are expressed at similar levels (2, 3). It is therefore likely that these two genes are controlled by similar transcriptional regulatory mechanisms.

The DNA sequences important for the expression of the rat insulin genes have been approximately determined. With insulin-CAT (chloramphenicol acetyltransferase) fusion genes, Walker et al. (32) showed that the −300 to +51 fragment of the rat insulin I gene is sufficient for tissue-specific expression in transfected insulinoma cells. Subsequently, Edlund et al. (7) demonstrated that this fragment can be divided into two distinct sequence elements, the promoter and the enhancer, each of which can direct tissue-specific expression. For the rat insulin II gene, Hanahan (13) demonstrated that 520 base pairs of 5′-flanking sequences linked to the simian virus 40 (SV40) T-antigen gene induced β-cell-specific tumor formation in transgenic mice. A similar fragment is sufficient for tissue-specific expression in cultured cells, as shown by Episkopou et al. (8). In addition to the positive elements, sequences at 2 to 4 kilobases (15) and 250 base pairs (22) upstream from the cap site inhibit expression of the rat insulin I gene and may help to repress insulin gene expression in inappropriate cell types.

To gain more insight into this complex regulation system, it is necessary to define the critical regulatory sequences in more detail and to study their interaction with *trans*-acting factor(s). Here we report the definition of a sequence element necessary for the expression of the rat insulin II gene and the identification of a protein binding to this region. Using extensively purified protein preparations, we determined that the binding protein is the same transcription factor that binds to the upstream promoter of the ovalbumin gene.

MATERIALS AND METHODS

Plasmids. The construction of the linker-scanning (LS) mutants will be described in detail elsewhere (D. T. Crowe and M.-J. Tsai, manuscript in preparation). Briefly, a series of 5′ and 3′ deletion mutants of the rat insulin II gene were constructed using either BAL 31 or exonuclease III-S1 nuclease, and the precise endpoints were determined by double-stranded sequencing with a ^{32}P-end-labeled sequencing primer (20). Appropriate pairs of 5′ and 3′ deletion mutants were then recombined to make LS mutants according to the method of Haltiner et al. (12). The resulting LS mutants (Fig. 1B) have the −448 to +49 sequence of the rat insulin II gene, in which small regions have been replaced by a *Sac*I linker (CGAGCTCG). These sequences were then linked to the bacterial CAT gene and the SV40 splice-polyadenylation signals.

Transfection. Transient transfection experiments were performed according to existing protocols (7) except that 1.5 × 10^6 HIT-T15 M2.2.2 cells were plated per 6-cm dish on the day before transfection. HIT cells were maintained in Dulbecco modified Eagle medium–12.5% donor equine serum–2.5% fetal calf serum–100 U of penicillin per ml–100 μg of streptomycin per ml, in the presence of 5% CO_2 at 37°C. Cells were harvested 48 to 60 h posttransfection and then assayed for CAT activity as previously described (11).

Band-shifting assays. The insulin promoter probe was the *Hae*III (−100)-to-*Rsa*I (+49) fragment of the rat insulin II gene, with a *Hin*dIII linker attached to the *Rsa*I site. It was end labeled at the *Hin*dIII site with Klenow enzyme (Boehringer Mannheim Biochemicals) and [α-^{32}P]dATP. The ovalbumin promoter probe was labeled similarly with [α-^{32}P]dCTP at a *Cla*I site which was attached to the −44 position of the −269 to −44 fragment of the chicken ovalbumin gene. The band-shifting assay was performed as

* Corresponding author.

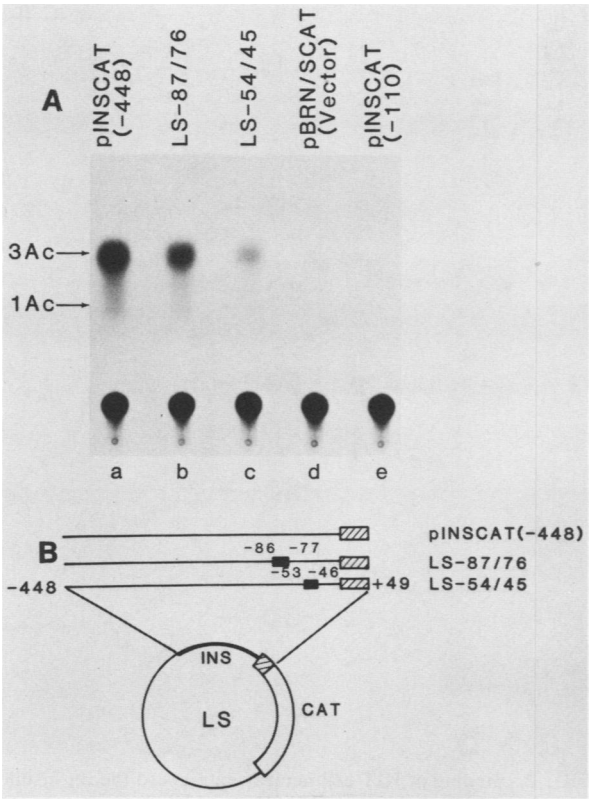

FIG. 1. Sequences between −54 and −45 are required for rat insulin II promoter activity. (A) CAT activities in HIT cells transfected with the following plasmids: lane a, the wild-type plasmid pINSCAT (−448); lane b, LS-87/76; lane c, LS-54/45; lane d, the vector with no insulin gene sequences, pBRN/SCAT; lane e, an enhancerless 5′-deletion mutant containing the −110 to +49 region of the rat insulin II gene. 3Ac, 3-Acetylated form of chloramphenicol; 1Ac, 1-acetylated form of chloramphenicol. (B) Structure of the LS mutants. The thick line and the hatched box are the 5′-flanking and coding regions of the rat insulin II gene, respectively. The open box has the CAT gene and SV40 splice-polyadenylation signals. The thin line is a vector sequence which contains the *Acc*I (2249)-to-*Cla*I (25) fragment of pBR322. The filled boxes represent regions replaced by *Sac*I linkers in the LS mutants.

previously described (27) except that poly(dI-dC) · poly(dI-dC) (Pharmacia) was used as a nonspecific competitor instead of pBR322/*Hin*fI and no $MgCl_2$ was included in the reaction mixture.

DNase I-footprinting assays. The insulin promoter probe used in DNase I-footprinting assays was identical to that used in band-shifting assays unless otherwise indicated. DNase I footprinting was carried out as previously described (24), except that reaction conditions identical to those of the band-shifting assay were used, and the concentrations of poly(dI-dC) · poly(dI-dC), NaCl, and DNase I were experimentally determined, as indicated in the figure legends.

Nuclear extracts. Nuclear extracts of HeLa cells were made as previously described (24). Nuclear extracts of HIT cells were prepared as follows. HIT cells grown in roller bottles were treated with phosphate-buffered saline (PBS) containing 0.1 M EDTA at room temperature and then harvested, using a rubber policeman. The cells were centrifuged at 1,600 × *g* for 5 min at 4°C, washed with PBS, and recentrifuged. The packed cell volume (PCV) was measured, and the cells were stored at −70°C. To prepare nuclear extracts, frozen cells were thawed on ice, suspended in 2 PCV of PBS containing 1 mM dithiothreitol (DTT) and 0.5 mM phenylmethylsulfonyl fluoride (PMSF), and centrifuged in a Beckman JA20 rotor at 4,400 × *g* for 5 min at 2°C. The pellets were suspended in 5 PCV of buffer A (10 mM HEPES [*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid] [pH 7.9], 1.5 mM $MgCl_2$, 10 mM NaCl, 0.1 mM EDTA, 2 mM DTT, and 0.5 mM PMSF), incubated on ice for 10 min, and centrifuged in a JA20 rotor at 4,400 × *g* for 5 min at 2°C. The pellets were suspended in 2 PCV of buffer A containing 10 μg each of aprotinin (Boehringer Mannheim Biochemicals) and leupeptin (Boehringer Mannheim Biochemicals) per ml, homogenized with the Dounce homogenizer (pestle B) for 10 strokes, and centrifuged in a JA20 rotor at 12,000 × *g* for 10 min at 2°C. The pellets were suspended in 0.5 PCV of buffer T (20 mM HEPES [pH 7.9], 0.1 mM EDTA, 100 mM NaCl, 20% glycerol, 2 mM DTT, and 0.5 mM PMSF), adjusted to 0.4 M NaCl with buffer T containing 4 M NaCl, and stirred for 30 min at 4°C in the presence of 10 μg of aprotinin and leupeptin per ml. The mixture was then centrifuged in a JA20 rotor at 39,000 × *g* for 30 min at 2°C. The supernatant was dialyzed against buffer T overnight at 4°C and centrifuged in a JA20 rotor at 27,000 × *g* for 20 min at 2°C. The supernatant (nuclear extract) was aliquoted and stored at −70°C.

Purification of the COUP transcription factor. Purification of the COUP (chicken ovalbumin upstream promoter) transcription factor was performed as previously described (33). Briefly, nuclear extracts of HeLa cells were fractionated by DEAE, phosphocellulose, Sephacryl S300, and heparin-Sepharose column chromatography. The fractions containing COUP transcription factor were pooled and loaded onto a COUP sequence-specific affinity column, and the flow-through fraction was collected. The column was then washed with a buffer containing 0.3 M NaCl, and three 1-ml fractions were collected. The COUP transcription factor was subsequently eluted with a buffer containing 0.6 M NaCl.

Renaturation of the COUP transcription factor. Renaturation of the COUP transcription factor was performed as previously described (33).

RESULTS

RIPE is essential for rat insulin II gene expression. To delimit the sequences essential for the expression of the rat insulin II gene, we constructed a series of LS mutants in the 5′-flanking region of an insulin II-CAT fusion gene. The mutations in two of the LS mutants reside in the promoter region (Fig. 1), where the sequences from −86 to −77 and −53 to −46 were replaced with a *Sac*I linker. After these two LS mutants, as well as the wild-type and negative controls (−448 and pBRN/SCAT, respectively), were transiently transfected into HIT cells, CAT enzyme activity was assayed. Quantitation of four independent transfections showed reproducibly that LS-87/76 yielded 50% of the CAT activity of the wild type, whereas LS-54/45 had only 15% of the CAT activity seen in the wild-type control. Since other LS mutants in the enhancer region display high levels of CAT activity, we conclude that the presence of the *Sac*I linker by itself is not deleterious to expression. These experiments have therefore defined at least one important promoter element which we call RIPE (rat insulin promoter element), located in the −54 to −45 region, that is important for the in vivo expression of the rat insulin II gene.

RIPE-binding factors. It is likely that RIPE exerts its function through interaction with *trans*-acting factor(s). To examine this possibility, we performed band-shifting assays,

using nuclear extracts from HIT cells and a probe harboring the −100 to +49 region of the insulin II gene. Multiple shifted bands were formed, indicating the presence of one or more binding factors (Fig. 2A). The binding site was then determined by DNase I-footprinting assays to be −60 to −40 (Fig. 2B), which fully covers the region previously identified by transfection experiments as being important for expression in vivo. Therefore, this binding activity is likely to be an important factor for expression of the rat insulin II gene.

HeLa cell nuclear extracts were also used in binding assays to examine whether the RIPE-binding factor is present in non-insulin-producing cells. These extracts yielded a band-shifting pattern similar to that of HIT cells (Fig. 3A), suggesting that HeLa cell factor(s) can bind to the insulin promoter and that HIT and HeLa cells may share some common factor(s). Furthermore, a fraction of HeLa nuclear extract purified by DEAE, phosphocellulose, and S300 column chromatography footprints the −60 to −40 region of the insulin promoter (Fig. 3B), indicating that a RIPE-binding activity is present. These results demonstrate that RIPE-binding factors are not restricted to insulin-producing cells. In addition to HeLa cells, similar binding activities were also found in chicken oviduct, rat liver, and pancreatic extracts (data not shown).

In band-shifting assays, the S300 fraction from HeLa cells forms a prominent complex which migrates to the same position as the no. 2 band in both HIT and HeLa cell extracts (Fig. 3A). Taken together with the footprinting analyses, this suggests that the RIPE-binding factor is responsible for the formation of the no. 2 complex. This conclusion is supported by indirect footprinting analysis. The insulin promoter probe was incubated with the HeLa S300 fraction under conditions identical to those of the band-shifting assay and then treated with DNase I before being loaded onto the gel. The no. 2 band was excised after electrophoresis, and the DNA contained in this band was eluted and analyzed in a sequencing gel. A footprint identical to that shown in Fig. 3B was observed (data not shown), indicating that the binding factor in the no. 2 complex is the RIPE-binding factor.

We reasoned that if binding of the RIPE factor to the −60 to −40 region is important for transcription of the insulin gene, then mutations in this region which disrupt gene expression should also affect the binding of the RIPE factor. To test this hypothesis, we carried out band-shifting assays, using the promoter fragments of the wild-type plasmid, LS-87/76 and LS-54/45. The promoter fragments from the wild-type plasmid and LS-87/76, which showed normal levels of activity in transfection experiments, bound normally to the RIPE-binding factor in HIT (Fig. 4, lanes b and h) and HeLa cells (Fig. 4, lanes c and i). In contrast, the promoter fragment from LS-54/45, which displayed markedly reduced activity in transfection experiments, barely bound the RIPE-binding factor in HIT or HeLa cells (Fig. 4, lanes e and f, respectively). Thus, the binding activity in vitro correlates well with promoter activity in vivo, supporting the conclusion that the RIPE-binding factor plays a functional role in the expression of the insulin gene. These results also further support the notion that the no. 2 band in band-shifting experiments results from protein binding to the RIPE sequence.

The HeLa cell RIPE-binding factor copurified with the COUP transcription factor. The RIPE-binding factor was further purified from HeLa cells. Surprisingly, it copurified with the COUP transcription factor, which is required for efficient transcription of the ovalbumin gene in vitro (27). The COUP transcription factor has been purified more than

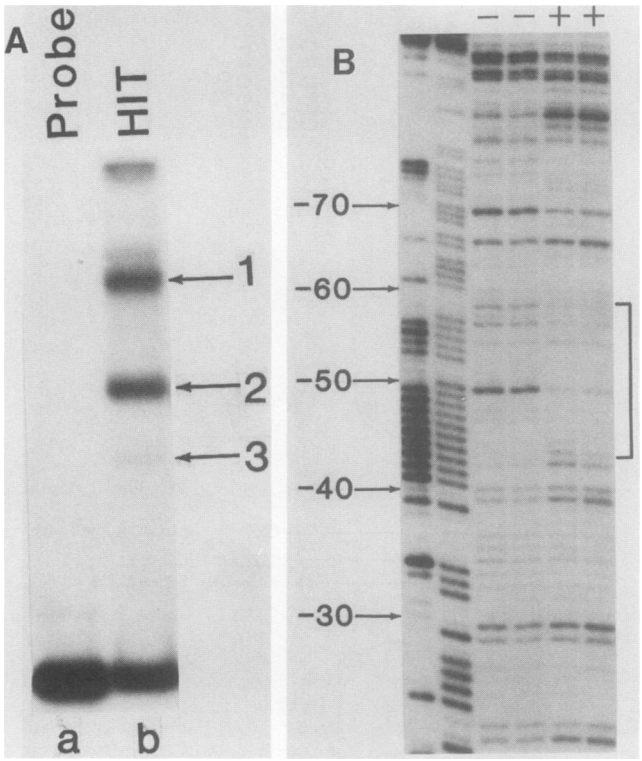

FIG. 2. Binding of HIT cell nuclear extracts to the rat insulin II promoter. (A) Band-shifting assay. An end-labeled promoter fragment (−100 to +49; 0.2 ng) was used in each reaction. Lane a, Buffer control; lane b, HIT cell nuclear extract containing 5 μg of protein; 4 μg of poly(dI-dC) · poly(dI-dC) was used as a nonspecific competitor. (B) DNase I footprinting assay. Left two lanes, G and G+A sequence ladders, respectively; −, buffer controls; +, HIT cell nuclear extract containing 50 μg of protein. The DNase I concentrations for lanes 3 to 6 from the left are 1, 2, 6, and 9 μg/ml, respectively. 0.2 ng of probe and 50 ng of nonspecific competitor, poly(dI-dC) · poly(dI-dC), were used in each reaction. The footprint is indicated by a bracket.

100,000-fold by a protocol that includes sequence-specific DNA affinity chromatography (33). In each purification step, the protein preparation pooled for COUP transcription factor activity also binds to the insulin promoter. Figure 5 shows the results of binding assays with the COUP affinity column fractions. The flowthrough and the 0.3 M NaCl fractions, which contain many DNA-binding proteins, do not bind to the insulin promoter, while the 0.6 M NaCl fraction containing the COUP transcription factor shows the RIPE-binding activity.

This copurification was unexpected, since the COUP transcription factor binds to the ovalbumin promoter at −90 to −70, where the sequence (5′-TATGGTGTCAAAGGT CAAACTT-3′) is quite different from the RIPE factor-binding site (5′-CCAGGGGTCAGGGGGGGGGTGCTT-3′). Moreover, the ovalbumin promoter-binding site has a 36% GC content, which is far less than the 75% of the insulin promoter-binding site. Therefore, we carried out competition and renaturation experiments to determine whether the two binding proteins are identical or if the insulin promoter-binding factor is a contaminant in the COUP factor preparation.

The HeLa cell RIPE-binding factor is the same as the COUP transcription factor. In the competition experiments, a radioactively labeled insulin promoter fragment was used as a

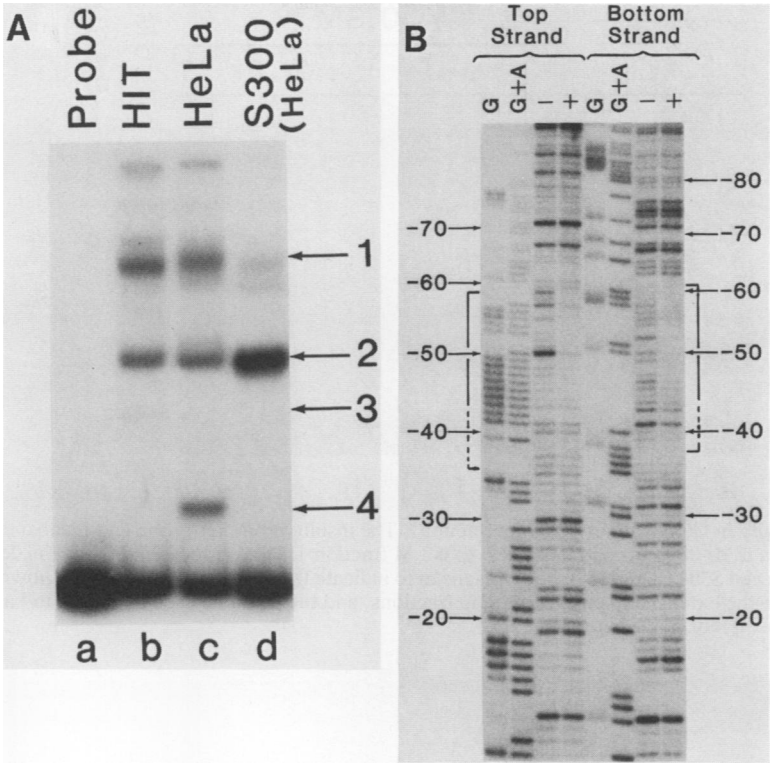

FIG. 3. Binding of various protein fractions to the insulin promoter. (A) Band-shifting assay. Lane a, Buffer control; lane b, HIT cell nuclear extract (5 μg of protein); lane c, HeLa cell nuclear extract (5 μg of protein); lane d, HeLa S300 fraction (1.8 μg of protein). Poly(dI-dC) · poly(dI-dC) (4 μg) was used in lanes b to d. (B) DNase I footprinting of HeLa cell factors. G and G+A represent the G and G+A sequence ladders, respectively. −, Buffer control; +, HeLa S300 fraction (9 μg of protein). The top strand probe was the insulin promoter probe described in Materials and Methods. The bottom strand probe contained the same DNA fragment as the top strand except that it was dephosphorylated at the *Hin*dIII site and then labeled with [γ-^{32}P]ATP and T4 DNA kinase. Probe (0.4 ng) and 1 μg of poly(dI-dC) · poly(dI-dC) were used in each reaction. The nucleotide numbers relative to the transcription start are shown on both sides. The footprinted regions are shown by brackets, with the dotted lines indicating weak protection.

probe for the binding of the RIPE factor. Incubation of the probe with excess amounts of unlabeled ovalbumin or insulin promoter fragment before the addition of the HeLa S300 fraction drastically reduced complex formation (Fig. 6A), indicating that the ovalbumin promoter competes efficiently with the insulin promoter for the RIPE-binding factor. Similarly, a converse competition experiment was performed with a labeled ovalbumin promoter probe for the COUP transcription factor. The insulin promoter competed with the ovalbumin promoter for binding of the COUP transcription factor (Fig. 6B). A control ovalbumin fragment with the COUP sequence deleted [OV (−103 −786; Fig. 6B)] does not compete, implicating the COUP-specific nature of the binding factor. Therefore, the COUP transcription factor also binds to both promoters. These results are consistent with the possibility that the RIPE-binding factor is the same as the COUP transcription factor. In addition, the results indicate that the factor(s) has a higher affinity for the ovalbumin promoter than for the insulin promoter, since it requires less ovalbumin competitor DNA to achieve a given level of competition than the insulin promoter does.

To examine the binding specificity of the factor(s) in more detail, we used synthetic oligonucleotides containing wild-type or mutated COUP-binding sequences as competitors in binding assays. The wild-type oligonucleotide is a 24-mer representing the −91 to −68 region of the ovalbumin gene, and the mutated oligonucleotide has two G-to-C base substitutions at −81 and −74 on the bottom strand, both of

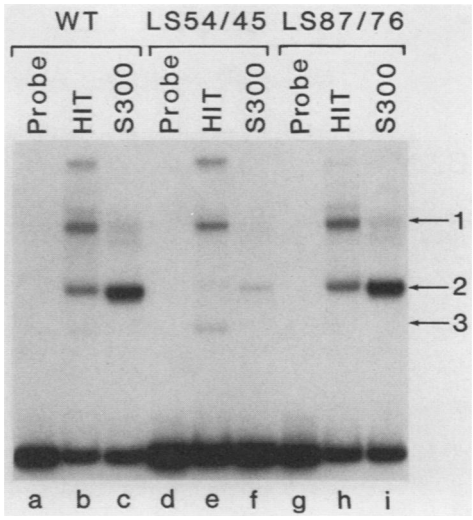

FIG. 4. Band-shifting assay with insulin promoter fragments from wild-type plasmid, LS-54/45, and LS-87/76. The probes used are the following: lanes a to c, wild type; lanes d to f, LS-54/45; lanes g to i, LS-87/76. All probes contain the *Rsa*I (+49)-to-*Hae*III (−100) fragments with *Hin*dIII linkers at +49. The protein preparations are the following: lanes a, d, and g, buffer control; lanes b, e, and h, HIT cell nuclear extract (5 μg of protein); lanes c, f, and i, HeLa S300 fractions (1.5 μg of protein). Poly(dI-dC) · poly(dI-dC) (4 μg) was present in each reaction.

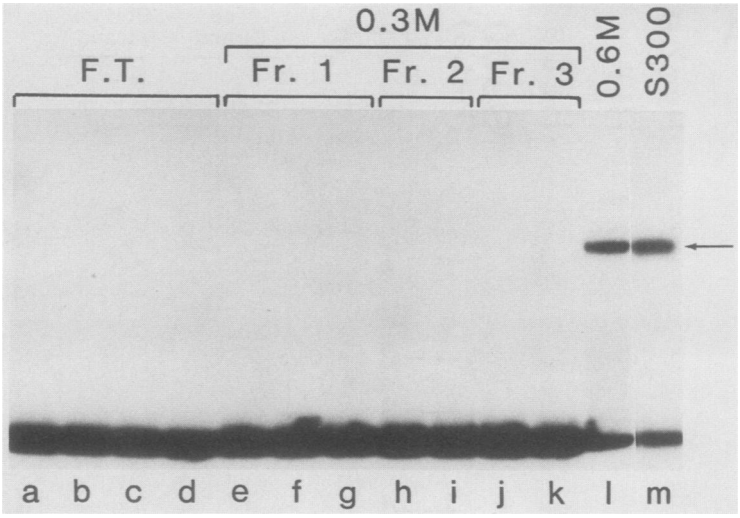

FIG. 5. RIPE-binding activity in COUP affinity column fractions. The insulin promoter probe was incubated with various fractions of the COUP affinity column: lanes a to d, flowthrough; lanes e to g, 0.3 M fraction 1; lanes h to i, 0.3 M fraction 2; lanes j and k, 0.3 M fraction 3; lane l, 0.6 M fraction. HeLa cell S300 fraction was used in lane m to indicate the position of band no. 2 (shown by the arrow). Bovine serum albumin was added to stabilize the extensively purified protein fractions, and the amounts used in lanes a to l are 1, 2, 4, 8, 0.5, 1.5, 2.5, 0.5, 1.5, 0.5, 1.5, and 1.5 μg, respectively. Fr, Fraction.

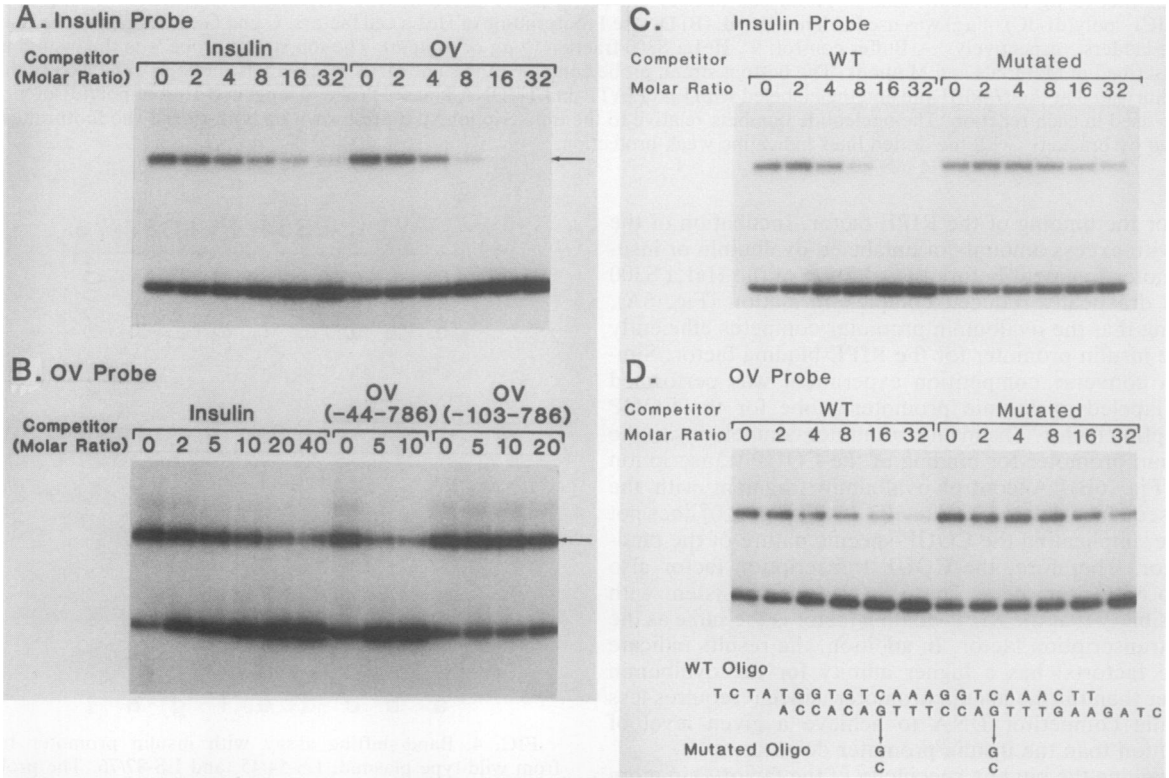

FIG. 6. Binding competition experiments. Each lane contained 0.2 ng of probe, HeLa S300 fraction (1 μg of protein), and 3 μg of poly(dI-dC) · poly(dI-dC). The molar ratio of competitor to probe is shown for each lane. Arrows indicate the bands of interest. The insulin and ovalbumin (OV) probes are described in Materials and Methods. (A and B) The insulin competitor is a *Hin*dIII-*Bgl*I fragment containing the −103 to +49 insulin gene fragment with a *Hin*dIII linker at +49. The ovalbumin competitor in panel A is the −44 to −753 fragment of the ovalbumin gene, with a *Cla*I linker at −44. The ovalbumin competitors in panel B are the ovalbumin fragments shown in parentheses. (C and D) WT, Wild-type ovalbumin oligonucleotide (shown on bottom). The mutated oligonucleotide has two base substitutions.

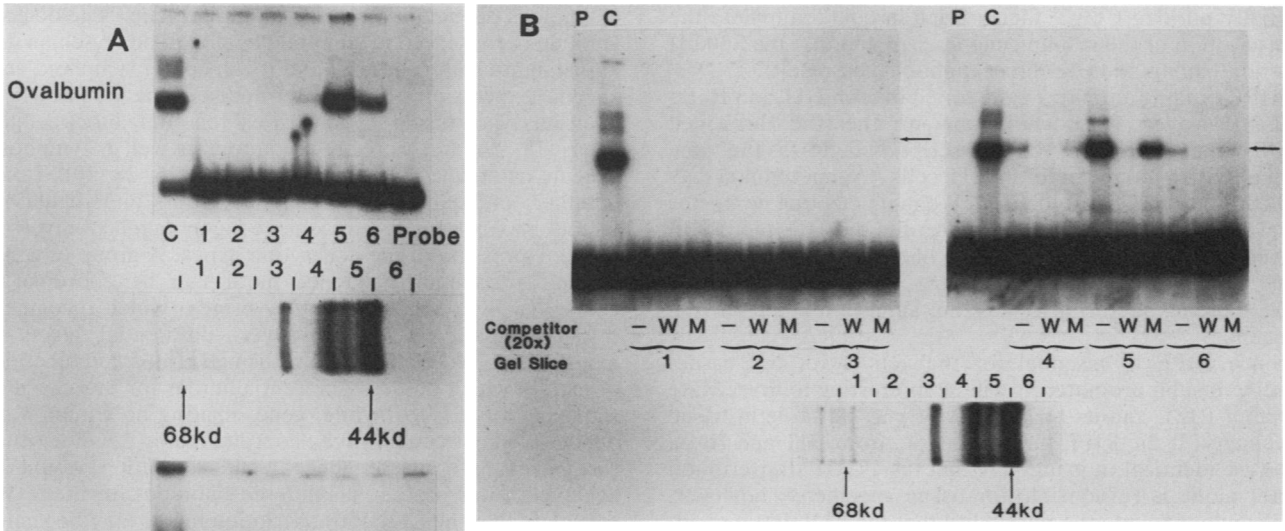

FIG. 7. Binding assays with renatured COUP transcription factor. (A) The affinity column-purified COUP transcription factor was resolved by sodium dodecyl sulfate-polyacrylamide gel electrophoresis; the silver-stained gel is shown in the middle. The gel was cut into six slices indicated by the numbers and vertical bars. A polypeptide(s) was renatured from each slice and used in band-shifting assays with the ovalbumin or the insulin promoter probe. The protein preparation used in each lane is indicated. C, Controls containing the HeLa S300 fraction; Probe, probe only. A 7- or 10-μl portion of renatured protein was used in each reaction for the ovalbumin and the insulin probe, respectively. (B) Binding competition assays using renatured COUP transcription factor, insulin promoter probe, and COUP oligonucleotides. P, Probe only; C, controls containing HeLa S300 fractions. Renatured protein from gel slices 1 to 6 is indicated for each lane. −, No competitor; W, wild-type COUP oligonucleotide competitor; M, mutated COUP oligonucleotide (Fig. 6). The molar ratio of competitor to probe is 20. The arrows indicate the shifted bands of interest. kd, Kilodaltons.

which are purine contacts for the COUP transcription factor (31). As expected, the wild-type oligonucleotide competes well with the ovalbumin promoter, while the mutated oligonucleotide does not (Fig. 6D). Similarly, the binding of the RIPE-binding factor to the insulin promoter is also competed for by the wild-type but not by the mutated COUP oligonucleotide (Fig. 6C). Again these results suggest that the RIPE-binding factor is the same as the COUP transcription factor.

The identity of the binding protein was confirmed by renaturation experiments. The affinity column-purified COUP transcription factor was resolved on sodium dodecyl sulfate-polyacrylamide gel, and 5-mm slices were excised. The polypeptide(s) in each slice was renatured and used in band-shifting assays. Renatured polypeptides from slices 3 to 6 bind to the ovalbumin promoter and the insulin promoter in a similar manner (Fig. 7A). Furthermore, the binding of the renatured polypeptides in each slice to the insulin promoter is competed for by the wild-type COUP oligonucleotide but not by the mutated oligonucleotide (Fig. 7B). Taken together, these results clearly indicate that the RIPE-binding protein in HeLa cells is the COUP transcription factor which binds to the ovalbumin upstream promoter.

DISCUSSION

Evidence has been presented that the COUP transcription factor, a binding factor of the ovalbumin promoter, also binds to the rat insulin promoter element (RIPE) which is important for rat insulin II promoter function. When part of the binding region is mutated, both the in vitro binding and the in vivo promoter activities are drastically reduced, suggesting that the COUP transcription factor may be part of the transcriptional machinery for the rat insulin II gene. Previously, our laboratory has shown that both the COUP transcription factor and S300-II, a non-DNA-binding factor isolated from HeLa cells, are required for efficient transcription of the ovalbumin gene (27). S300-II appears to function by interacting with the COUP transcription factor, resulting in the stabilization of the DNA-COUP factor complex (31). It will be interesting to see whether the same mechanism also applies to the insulin gene. Preliminary data suggest that a

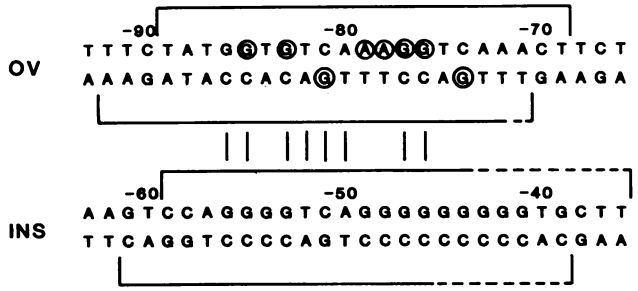

FIG. 8. Alignment of the COUP transcription factor-binding sequences in the ovalbumin (OV) and the insulin (INS) genes. The brackets indicate the footprinted regions; the broken lines represent weak protection. Purine contacts in the ovalbumin gene are shown by circles. Common nucleotides are indicated by vertical bars between the ovalbumin and insulin sequences.

partially purified COUP factor fraction could stimulate the transcription of the insulin gene in vitro and that the S300-II fraction had an additive effect (unpublished result).

RIPE-binding activities were found in both HIT and HeLa cells. However, we do not know whether the HeLa cell RIPE-binding factor (COUP transcription factor) is the same as the RIPE-binding factor in HIT cells. Competition assays indicate that the wild-type COUP oligonucleotide competes efficiently with the insulin promoter for the HIT cell RIPE factor, while the mutated COUP oligonucleotide does not (data not shown). This result suggests that the HIT cell RIPE-binding factor also has an affinity for the COUP sequence.

Since RIPE is essential for the activity of the tissue-specific insulin promoter, it will be interesting to investigate whether RIPE and its binding factor play any role in tissue specificity. If the RIPE-binding factors from HIT and HeLa cells are identical, it is unlikely that the COUP transcription factor alone is responsible for tissue specificity; however, one cannot exclude the possibility that the COUP transcription factor interacts with other tissue-specific or nonspecific factor(s) to mediate tissue specificity. If the RIPE-binding factor in HIT cells is either completely different or modified differently from the COUP factor, the RIPE-binding factor in HIT cells will be a potential candidate for directing tissue specificity.

Alignment of the COUP transcription factor-binding sequences of the insulin and ovalbumin promoters (Fig. 8) reveals eight common nucleotides. Of the eight purine contacts in the ovalbumin promoter (31) which may be crucial for its binding to the COUP transcription factor, five are shared by the insulin gene. Several possibilities may explain how one factor can bind to such diverse sequences. First, the eight common nucleotides may be sufficient for binding. Alternatively, the two different sequences may form a common structure which is recognized by the factor. Finally, the factor may possess two binding domains which have different specificities for the two sequences. We are presently unable to distinguish among these possibilities.

The renatured COUP polypeptides bind to the ovalbumin and insulin promoters in a similar manner (Fig. 7A). Thus, polypeptides from slice 3 bind weakly to both probes and shift them to a slightly higher position. Polypeptides from slice 5 bind strongly, while those from slices 4 and 6 have intermediate binding activities. However, slices 4 and 6 showed different binding preferences for the two promoters; slice 6 has a higher affinity for the ovalbumin promoter, whereas slice 4 favors the insulin promoter. This suggests that the COUP transcription factor may exist in heterogeneous forms, and each form may have different affinities for the two sequences. The heterogeneity may arise from a gene family, differential splicing or translational initiation and termination, posttranslational modification, or proteolytic degradation.

The fact that the COUP transcription factor binds to two seemingly unrelated genes through two rather different sequences is intriguing. One factor binding to two different sequences is not unprecedented; the yeast protein HAP1 was shown to bind to the upstream activation sites of both CYC1 and CYC7 genes, and the binding sites share only seven nucleotides scattered throughout the 23-base-pair footprint (25). In contrast, the same consensus sequence can also be recognized by different proteins. For instance, at least three different CCAAT-binding proteins have been found, and each of them has different affinities for the CCAAT-containing regions of various genes (4). Another example is the octanucleotide-binding proteins. The octanucleotide is conserved in all immunoglobulin heavy-chain and kappa light-chain genes (9), and it can confer lymphoid-cell-specific expression upon a heterologous gene (5, 34). However, it is also present in, and critical for, some other cellular genes (18, 29), and there are ubiquitous as well as lymphoid-specific octanucleotide-binding proteins (16, 28, 30). Taken together, multiple related sequence motifs and regulatory factors may be required to achieve higher levels of diversity in the expression of the eucaryotic genes. A group of genes sharing a common sequence motif can be coordinately controlled by a single regulatory protein which recognizes this common element. Moreover, differential levels of expression could be obtained by using various combinations of factors which possess different affinities for this common sequence motif. Meanwhile, genes sharing no common sequence element could also be coordinated by a factor which recognizes different sequences, and regulation would be achieved if this protein has different affinities for them. We speculate that the COUP transcription factor may be such a regulatory factor.

ACKNOWLEDGMENTS

We thank P. Lomedico of Hoffmann-La Roche Inc. for the rat insulin II gene clone, W. J. Rutter of the University of California at San Francisco for the HIT-T15 M2.2.2 cells, J. Codina-Salada for synthesizing oligonucleotides, and G. Long for excellent technical assistance. We also like to thank the members of the Tsai laboratory for helpful discussion and C. McCarthy for the preparation of the manuscript.

This work was supported by Public Health Service grant HD-17379 from the National Institutes of Health and a feasibility study grant from Baylor Diabetes Center.

LITERATURE CITED

1. **Chodosh, L. A., R. W. Carthew, and P. A. Sharp.** 1986. A single polypeptide possesses the binding and transcription activities of the adenovirus major late transcription factor. Mol. Cell. Biol. **6**:4723–4733.
2. **Clark, J. L., and D. F. Steiner.** 1969. Insulin biosynthesis in the rat: demonstration of two proinsulins. Proc. Natl. Acad. Sci. USA **62**:278–285.
3. **Cordell, B., D. Diamond, S. Smith, J. Pünter, H. H. Schöne, and H. M. Goodman.** 1982. Disproportionate expression of the two nonallelic rat insulin genes in a pancreatic tumor is due to translational control. Cell **31**:531–542.
4. **Dorn, A., J. Bollekens, A. Staub, C. Benoist, and D. Mathis.** 1987. A multiplicity of CCAAT box-binding proteins. Cell **50**:863–872.
5. **Dreyfus, M., N. Doyen, and F. Rougeon.** 1987. The conserved decanucleotide from the immunoglobulin heavy chain promoter induces a very high transcriptional activity in B-cell when introduced into an heterologous promoter. EMBO J. **6**:1685–1690.
6. **Dynan, W. S., and R. Tjian.** 1985. Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. Nature (London) **316**:774–778.
7. **Edlund, T., M. D. Walker, P. J. Barr, and W. J. Rutter.** 1987. Cell-specific expression of the rat insulin gene: evidence for role of two distinct 5′-flanking elements. Science **230**:912–916.
8. **Episkopou, V., A. J. M. Murphy, and A. Efstratiadis.** 1984. Cell-specified expression of a selectable hybrid gene. Proc. Natl. Acad. Sci. USA **81**:4657–4661.
9. **Falkner, F. G., and H. G. Zachau.** 1984. Correct transcription of an immunoglobulin κ gene requires an upstream fragment containing conserved sequence elements. Nature (London) **310**: 71–74.
10. **Gluzman, Y.** 1985. Eukaryotic transcription—the role of cis- and trans-acting elements in initiation. Current communications in molecular biology. Cold Spring Harbor Laboratory, Cold

Spring Harbor, New York.
11. **Gorman, C. M., L. F. Moffat, and B. M. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
12. **Haltiner, M., T. Kempe, and R. Tjian.** 1985. A novel strategy for constructing clustered point mutations. Nucleic Acids Res. **13**:1015–1224.
13. **Hanahan, D.** 1985. Heritable formation of pancreatic β-cell tumors in transgenic mice expressing recombinant insulin/simian virus 40 oncogenes. Nature (London) **315**:115–122.
14. **Jacob, F., and J. Monod.** 1961. Genetic regulatory mechanisms in the synthesis of proteins. J. Mol. Biol. **3**:318–356.
15. **Laimins, L., M. Holmgren-König, and G. Khoury.** 1986. Transcriptional "silencer" element in rat repetitive sequences associated with the rat insulin I gene locus. Proc. Natl. Acad. Sci. USA **83**:3151–3155.
16. **Landolfi, N. F., J. D. Capra, and P. W. Tucker.** 1986. Interaction of cell-type-specific nuclear proteins with immunoglobulin V_H promoter region sequences. Nature (London) **323**:548–551.
17. **Lomedico, P., N. Rosenthal, A. Efstratiadis, W. Gilbert, R. Kolodner, and R. Tizard.** 1979. The structure and evolution of the two nonallelic rat preproinsulin genes. Cell **18**:545–558.
18. **Mangin, M., M. Ares, and A. M. Weiner.** 1986. Human U2 small nuclear RNA genes contain an upstream enhancer. EMBO J. **5**:987–996.
19. **Maniatis, T., S. Goodbourn, and J. A. Fischer.** 1987. Regulation of inducible and tissue-specific gene expression. Science **236**: 1237–1245.
20. **McGraw, R. A.** 1984. Dideoxy DNA sequencing with end-labeled oligonucleotide primers. Anal. Biochem. **143**:298–303.
21. **McKnight, S., and R. Tjian.** 1986. Transcriptional selectivity of viral genes in mammalian cells. Cell **46**:795–805.
22. **Nir, U., M. D. Walker, and W. J. Rutter.** 1986. Regulation of rat insulin I gene expression: evidence for negative regulation in nonpancreatic cells. Proc. Natl. Acad. Sci. USA **83**:3180–3184.
23. **Parker, C. S., and J. Topol.** 1984. A Drosophila RNA polymerase II transcription factor binds to the regulatory site of an hsp 70 gene. Cell **37**:273–283.
24. **Pastorcic, M., H. Wang, A. Elbrecht, S. Y. Tsai, M.-J. Tsai, and B. W. O'Malley.** 1986. Control of transcription initiation in vitro requires binding of a transcription factor to the distal promoter of the ovalbumin gene. Mol. Cell. Biol. **6**:2784–2791.
25. **Pfeifer, K., T. Prezant, and L. Guarente.** 1987. Yeast HAP1 activator binds to two upstream activation sites of different sequence. Cell **49**:19–27.
26. **Reinberg, D., M. Horikoshi, and R. G. Roeder.** 1987. Factors involved in specific transcription in mammalian RNA polymerase II. J. Biol. Chem. **262**:3322–3330.
27. **Sagami, I., S. Y. Tsai, H. Wang, M.-J. Tsai, and B. W. O'Malley.** 1986. Identification of two factors required for transcription of the ovalbumin gene. Mol. Cell. Biol. **6**:4259–4267.
28. **Sen, R., and D. Baltimore.** 1986. Multiple nuclear factors interact with the immunoglobulin enhancer sequences. Cell **46**:705–716.
29. **Sive, H. L., N. Heintz, and R. G. Roeder.** 1986. Multiple sequence elements are required for maximal in vitro transcription of a human histone H2B gene. Mol. Cell. Biol. **6**:3329–3340.
30. **Staudt, L. M., H. Singh, R. Sen, T. Wirth, P. A. Sharp, and D. Baltimore.** 1986. A lymphoid-specific protein binding to the octamer motif of immunoglobulin genes. Nature (London) **323**: 640–643.
31. **Tsai, S. Y., I. Sagami, H. Wang, M.-J. Tsai, and B. W. O'Malley.** 1987. Interactions between a DNA-binding transcription factor (COUP) and a non-DNA binding factor (S300-II). Cell **50**:701–709.
32. **Walker, M. D., T. Edlund, A. M. Boulet, and W. J. Rutter.** 1983. Cell-specific expression controlled by the 5'-flanking region of insulin and chymotrypsin genes. Nature (London) **306**:557–561.
33. **Wang, L.-H., S. Y. Tsai, I. Sagami, M.-J. Tsai, and B. W. O'Malley.** 1987. Purification and characterization of chicken ovalbumin upstream promoter transcription factor from HeLa cells. J. Biol. Chem. **262**:16080–16086.
34. **Wirth, T., L. Staudt, and D. Baltimore.** 1987. An octamer oligonucleotide upstream of a TATA motif is sufficient for lymphoid-specific promoter activity. Nature (London) **329**: 174–178.