MOLECULAR AND CELLULAR BIOLOGY, Apr. 1990, p. 1564–1572
0270-7306/90/041564-09$02.00/0
Copyright © 1990, American Society for Microbiology
Vol. 10, No. 4

Identification of a Pancreatic β-Cell Insulin Gene Transcription Factor That Binds to and Appears To Activate Cell-Type-Specific Expression: Its Possible Relationship to Other Cellular Factors That Bind to a Common Insulin Gene Sequence

JAMES WHELAN, SUSAN R. CORDLE, EVA HENDERSON, P. ANTHONY WEIL, AND ROLAND STEIN*

Department of Molecular Physiology and Biophysics, Vanderbilt University School of Medicine, Nashville, Tennessee 37232-0615

Received 29 September 1989/Accepted 21 December 1989

The insulin gene is expressed almost exclusively in pancreatic β-cells. Previous work in our laboratory has shown that pancreatic β-cell-specific expression of the rat insulin II gene is controlled by a number of positive and negative *cis*-acting DNA elements within the enhancer. We have shown that one element within the enhancer, located between nucleotides −100 and −91 (GCCATCTGCT; referred to as the insulin control element [ICE]) relative to the transcription start site, is controlled by both positive- and negative-acting cellular transcription factors. The positive-acting factor appears to be uniquely active in β-cells. To identify the nucleotides within the ICE that mediate positive cell-type-specific regulation, point mutations within this element were generated and assayed for their effects on expression. Base pairs −97, −94, −93, and −92 were found to be crucial for the activator function of this region, while mutations at base pairs −100, −96, and −91 had little or no effect on activity. The gel mobility shift assay was used to determine whether specific cellular factors associated directly with the ICE. Several specific protein-DNA complexes were detected in extracts prepared from insulin-producing and non-insulin-producing cells, including a complex unique to β-cell extracts. The ability of unlabeled wild-type and point mutant versions of the ICE to compete for binding to these cellular factors demonstrated that the β-cell-specific complex appears to contain the insulin gene activator protein(s). Interestingly, the adenovirus type 2 major late promoter upstream element (USE; GCCACGTGAC) also competed in the gel mobility shift assay for binding of cellular proteins to the ICE. These results suggested that the cellular factor that binds to the USE (i.e., USF) also interacts with the ICE. This was directly demonstrated by showing that ICE and USE sequences competed for the USF required for adenovirus type 2 major late promoter transcription in vitro and by showing that reticulocyte lysate-translated human USF products bound to the ICE. However, the USE sequences were unable to stimulate β-cell-type-specific activity in vivo. We discuss the possible relationship of these observations to positive and negative control mediated by the ICE.

In a multicellular organism, many genes are under developmental control and are transcribed in specific differentiated cell types. The mechanisms by which tissue-specific transcriptional control is achieved is currently the subject of extensive investigation. Cell-type- and tissue-specific regulation of eucaryotic gene transcription appears to be mediated by binding of regulatory proteins (*trans*-acting factors) to specific DNA sequences (*cis*-acting elements) usually, but not always, located in the 5′-flanking region of the gene (promoters-enhancers; for a review, see reference 27). Both enhancers and promoters consist of multiple short DNA-binding sites which are bound by *trans*-acting factors. There is increasing evidence that sequence-specific DNA-binding proteins confer transcriptional activity and specialized regulatory properties on enhancer-promoter elements (12, 17, 24, 40). Many *trans*-acting factors are constitutively active in various cell types (3, 4, 20, 39, 41), but others are unique to certain cell types (19, 23, 35, 45) or are inducible in response to extracellular stimuli that modulate gene expression (1, 24, 37, 42). The variety of binding sites and their combinatorial and spatial arrangements suggest that regulation of gene transcription results from complex interactions between multiple *trans*-acting factors. Many observations suggest, for example, that genes can be both positively and negatively regulated by interaction of multiple factors with overlapping binding sites, suggesting that promoter-enhancer geometry plays an important role in gene regulation (2, 14, 36, 49, 50). In addition, subtle differences in factor structure or small variations in the binding affinities of the factors might also be sufficient, in the appropriate context, to mediate specific regulatory events (18, 38, 44, 45). Together, these findings raise the possibility that different cell types contain related sets of DNA-binding proteins capable of interacting with overlapping DNA-regulatory elements in both stimulatory and inhibitory ways. Thus, gene regulation might result from small alterations in structure or relative concentrations of these interacting sets of binding proteins.

Our studies have been aimed at identifying the *cis*-acting elements and *trans*-acting factors involved in cell-specific control of rat insulin II gene transcription. Rats have two nonallelic insulin genes (termed I and II) which are almost equally transcribed in rat pancreas β-cells and are not expressed in other cell types (7, 8). Several lines of evidence, including transgenic mouse experiments, have demonstrated that β-cell-specific expression is regulated by 5′-flanking insulin gene enhancer and promoter sequence elements (10, 13, 16, 47). Analysis of the expression of various mutant rat insulin genes has demonstrated that cell type-specific

* Corresponding author.

expression is directed by the insulin gene enhancer, which lies between nucleotides −340 and −90 relative to the transcription start site.

Insulin gene enhancer-stimulated expression is mediated by multiple *cis*-acting elements (9, 10, 49). The activity of the enhancer appears to be controlled by both positive- and negative-acting cellular transcription factors. We have recently demonstrated that the activity of the *cis*-acting element located between −100 and −91 in the rat insulin II gene enhancer (referred to as the insulin control element [ICE]) is regulated by both of these cellular activities (49). Both the ICE and the corresponding elements in the rat insulin I gene, which are located between nucleotides −241 and −233 and −112 and −104, play a critical role in insulin enhancer expression (21). The importance of the ICE in β-cell-type-specific expression has recently been addressed. Multiple copies of this element were shown to be sufficient to selectively drive β-cell-specific expression of heterologous promoter expression vectors (22, 49). The ICE is found in the transcription units of all characterized insulin genes, and its core sequence motif (5′ −100 GCCATCTG −93 3′) is absolutely conserved (5). These results indicate that this element serves an important and general role in β-cell-specific expression of the insulin gene as a consequence of interaction with a β-cell-specific transcription factor.

To define the ICE and its role in cell-type-specific expression of the rat insulin II gene more clearly, we examined the effects of individual point mutations throughout this element. This analysis resulted in identification of a number of bases which are critical for activation of transcription by this element in β-cells. Using the gel mobility shift assay, we detected four cellular factor-DNA complexes which are formed with this sequence element. One of these four protein-DNA complexes was uniquely present in β-cell extracts. Formation of these complexes was differentially affected by the presence of wild-type or point mutant forms of the ICE. Pancreatic β-cell-type-specific expression appears to be mediated by an insulin activator factor(s) (IAF) found in the complex uniquely formed in extracts from these cells. Interestingly, one of the ubiquitously distributed factor-ICE complexes contained the specifically bound adenovirus type 2 upstream transcription factor (USF). However, additional in vitro and in vivo experiments demonstrated that USF does not function like a β-cell-specific activator. These results are discussed in regard to the possible involvement of the IAF, USF, and other ubiquitous factors in the generation of β-cell-specific transcription of the insulin gene.

MATERIALS AND METHODS

Cell lines. HeLa cells were grown in Spinner cultures in Dulbecco modified Eagle medium supplemented with 2.5% (vol/vol) fetal bovine serum, 2.5% (vol/vol) calf serum, 1 mM L-glutamine, and 1 mM nonessential amino acids. HeLa cells were grown in monolayer cultures in Dulbecco modified medium containing 10% (vol/vol) fetal bovine serum and 50 μg each of streptomycin and penicillin per ml. Insulin-producing cell lines HIT T-15 2.2.2 (HIT) and βTC-1 were grown in Dulbecco modified medium supplemented with 15% (vol/vol) horse serum, 2.5% (vol/vol) fetal bovine serum, and 50 μg each of streptomycin and penicillin per ml. HIT T-15 (10) and βTC-1 (11) cells are simian virus 40 (SV40) large-T-antigen-transformed hamster and mouse pancreatic β cells.

DNA transfection. Approximately 18 h before transfection, HeLa, βTC-1, and HIT cells were plated onto 100-mm^2 plates at densities of 0.5×10^6, 2×10^6, and 2×10^6, respectively. Transfection of plasmid DNA was performed by the calcium phosphate coprecipitation method under conditions described previously (49). At 4 h after addition of the calcium phosphate-DNA precipitate, the HIT and βTC-1 cells were exposed to 20% glycerol for 2 min. The amounts of plasmid DNA used in each transfection are indicated (see legend to Fig. 1). Each experiment was repeated several times with at least two different plasmid preparations.

Plasmids. The SV40 enhancer plasmids used in this study were prepared as follows. pSV−100 and pSV−90 contain the 2×72 SV40 enhancer insert from pβ2×72$^-$ (34) cloned into the *Hin*dIII site of the −100 and −90 rat insulin II chloramphenicol acetyltransferase (CAT) expression plasmids (49). The −100 to −91 wild-type and point mutation constructs were constructed by cloning into the *Xba*I-*Sal*I site, 5′ to the rat insulin II promoter (residues −90 to +8 [49]), oligonucleotides containing wild-type or single noncomplementary transversion mutations (A to C and T to G) in the rat insulin II sequences from −100 to −91. The sequences of all of the resulting plasmid constructs were confirmed by DNA sequencing. Insulin-CAT riboprobe expression plasmid pGEM-7 was prepared by subcloning rat insulin II-CAT sequences from −100 to +250 base pairs of the −100 rat insulin II CAT expression plasmid (49) as a *Hin*dIII-*Eco*RI fragment into pGEM-3Zf(+) (Promega Biotec, Madison, Wis.). Plasmids were prepared by being banded twice to equilibrium in cesium chloride gradients (26).

RNase mapping. The −100 to +250 coding strand transcription product was produced from *Hin*dIII-linearized pGEM-7 by using T7 RNA polymerase by a modification of the procedure described by Zinn et al. (52). One microgram of the linearized plasmid was transcribed in a 20-μl reaction with 50 μCi of [α-^{32}P]UTP (specific activity, 800 Ci/mmol) at a final UTP concentration of 12 μM. The labeled RNA (5×10^5 dpm) transcript was hybridized overnight at 45°C in 30 μl with 50 μg of total cellular RNA isolated by the guanidinium isothiocyanate method of Chirgwin et al. (6). Hybridization conditions, RNase treatment, and gel electrophoresis were conducted as described by Zinn et al. (52).

In vitro transcription assays. Cell-free transcription reactions contained 8 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid) (pH 7.9); 8% (vol/vol) glycerol; 7 mM $MgCl_2$; 2 to 2.5 μg of template DNA; 1 mg of bovine serum albumin per ml; 75 mM KCl; 80 μM EDTA; 80 μM EGTA [ethylene glycol-bis(β-aminoethylether)-*N,N,N′,N′*-tetraacetic acid], 80 μM dithiothreitol; and 500 μM each ATP, UTP, GTP, and CTP in a final assay volume of 25 μl. Approximately 100 μg of HeLa cell nuclear extract protein was incubated with template DNA and ribonucleotides at 30°C for 1 h. RNA was extracted with phenol-chloroform and analyzed by the primer extension method (28), and the primer extension products were visualized by autoradiography.

Extract preparation. HeLa and rat liver cell nuclear extracts were prepared as described by Shapiro et al. (43) and Gorski et al. (15), respectively. βTC-1 and HIT cell nuclear extracts were prepared as described by Shapiro et al. (43). Whole-cell extracts were prepared from βTC-1 and HIT cells as follows. Cells (approximately 1×10^9 to 2×10^9) were harvested and suspended in 20 ml of ice-cold buffer containing 400 mM ammonium sulfate, 20 mM HEPES (pH 7.9), 0.2 mM EDTA, 0.2 mM EGTA, 0.2 mM dithiothreitol, 0.2 mM benzamidine, and 5 mg each of leupeptin, pepstatin, antipain, and chymostatin per ml. The cells were homogenized

three times for 10 s each time with a Ultraturex brand homogenizer, and the homogenate was centrifuged at 100,000 × g for 60 min at 4°C. Solid ammonium sulfate was added, with stirring, to the supernatant (0.33 g/ml) and allowed to dissolve. The precipitate was collected by centrifugation at 100,000 × g for 20 min at 4°C and suspended in 2 ml of dialysis buffer (20 mM HEPES [pH 7.9], 0.2 mM EDTA, 0.2 mM EGTA, 0.2 mM dithiothreitol, 0.2 mM benzamidine, 10% [vol/vol] glycerol). Dialysis was performed for 4 h at 4°C. The extracts were then centrifuged at 10,000 × g for 5 min. Aliquots of the resulting supernatant were quick frozen on dry ice and stored at −70°C until use.

Gel shift assays. Double-stranded oligonucleotides containing the ICE (5'-TCTGGCCATCTGCTGATCCTCTG GCC-3') and USE (5'-TAGGTGTAGGCCACGTGACCGG GTGTTC-3') sequences were end labeled with ^{32}P-labeled deoxynucleotides and the Klenow fragment of *Escherichia coli* DNA polymerase I. Binding reactions contained 25 mM HEPES (pH 7.9); 5 mM dithiothreitol; 0.2 mM EDTA; 0.1 μg of poly(dI-dC) per μl; 0.1 μg of single-stranded DNA per μl; 9% (vol/vol) glycerol; 10 fmol of a ^{32}P-labeled double-stranded oligonucleotide (~5,000 dpm/fmol); 10 μg of extract protein; and 50, 125, or 200 mM KCl. Each 20-μl reaction was incubated at 4°C for 30 min. Samples were subjected to electrophoretic separation at 4°C on a 4% nondenaturing polyacrylamide gel at 200 V for 1.5 to 2 h by using the high-ionic-strength polyacrylamide gel electrophoresis conditions described by Carthew et al. (4). Following electrophoresis, the gels were dried and labeled DNA was localized by autoradiography.

In vitro transcription and translation. The plasmid containing the USF cDNA (kindly provided by P. Gregor and R. Roeder, manuscript in preparation) was linearized by digestion with *Xho*I. The in vitro transcription reaction was performed as described by Melton et al. (29), using 3 μg of the linearized plasmid. In vitro translation was performed as described in the Promega technical manual with 2 μg of in vitro-transcribed RNA. In the gel shift analyses, 2 μl of the in vitro translation reaction from a final volume of 50 μl was used in each reaction.

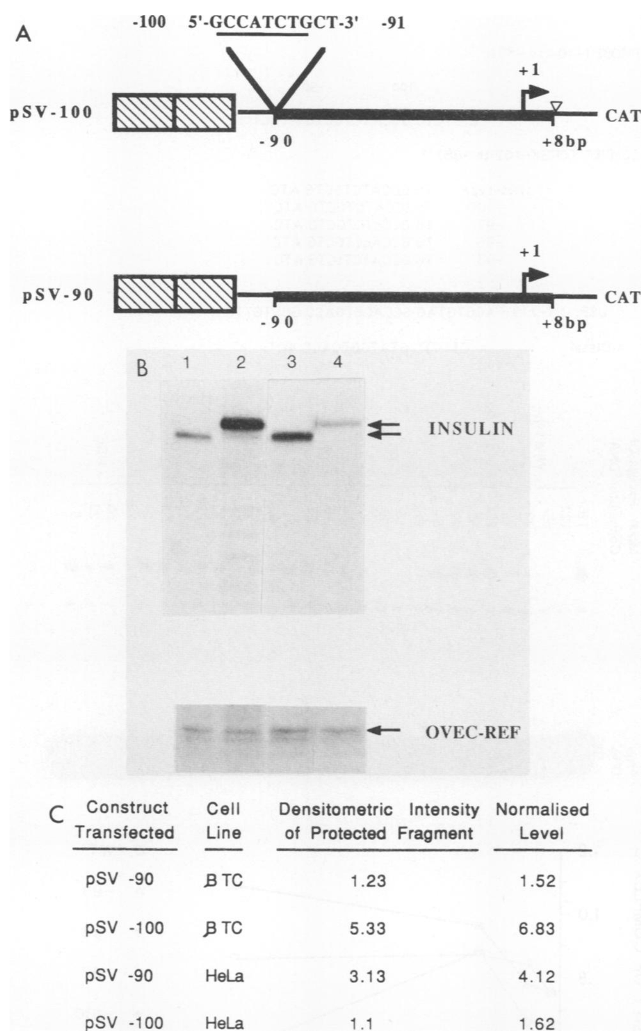
RESULTS

ICE activates transcription in a cell-type-specific manner when linked to a heterologous enhancer. Previously it was demonstrated that the ICE alone can direct β-cell-type-specific expression, although the activity mediated by it was significantly enhanced by upstream insulin enhancer sequence elements (10, 22, 49). In contrast, rat insulin II promoter activity does not appear to require β-cell-specific transcription factors. To test whether the ICE could also direct β-cell-type-specific expression from a heterologous enhancer element, the effect of the SV40 enhancer on rat insulin II gene promoter- and ICE-driven CAT expression was assayed in transfected βTC-1 (insulin-producing) and HeLa (non-insulin-producing) cells. CAT mRNAs from ICE expression plasmid pSV−100 and promoter expression plasmid pSV−90 were normalized to expression from a cotransfected SV40 enhancer-driven rabbit β-globin expression plasmid, OVEC-REF (48). The levels of the individual RNAs produced from these chimeric genes were quantitated by the RNase protection assay.

ICE expression plasmid pSV−100 was approximately fourfold more active than promoter expression plasmid pSV−90 in transfected insulin-producing cells (Fig. 1B; compare lanes 1 and 2). In contrast, the pSV−100 construct was approximately threefold less active than the pSV−90 construct in HeLa cells (Fig. 1B; compare lanes 3 and 4). The CAT enzyme activity levels from these transfected chimeras mirrored the relative steady-state mRNA levels shown in Fig. 1B (data not shown). The patterns of expression from these two plasmids are consistent with our previous observations (49). Thus, expression from the ICE appears to be mediated by a combination of positively (in insulin-producing cells) and negatively (in non-insulin-producing cells) acting transcription factors.

Identification of individual bases within the ICE that are required for activator function in vivo. To identify base pairs within the ICE that are important in cell-type-specific activation, each base pair within the element in pSV−100 was changed by noncomplementary transversion mutation (i.e., G to T and C to A). We reasoned that mutation of a base crucial for activator interaction would result in a decrease in activity from pSV−100 levels to the level found for promoter expression plasmid pSV−90. These plasmids were transfected into two different insulin-producing cell-lines, HIT and βTC-1. The CAT activity from the mutants was compared with the activity obtained from −100 to −91 wild-type element expression plasmid pSV−100 and rat insulin II gene promoter expression plasmid pSV−90, which were transfected in parallel. The effects of point mutations within the −100 to −91 element on β-cell expression are presented in Table 1. Mutations at −97, −94, −93, and −92 all reduced expression, while mutations at nucleotides −100, −96, and −91 had little effect on activity. Mutations at −99, −98, and −95 had an intermediate effect on activity. Equivalent results were obtained when expression of these constructs was examined by RNase protection analyses (data not shown). We conclude from these results that mutations within the −100 to −91 element at nucleotides −97, −94, −93, and −92 decrease the ability of a cellular activator to stimulate transcription in β-cells, presumably by decreasing the interaction of a positive-acting DNA-binding protein with this element.

Both cell-type-specific and ubiquitously distributed factors bind to the ICE. The gel mobility shift assay was used to determine whether cellular proteins interact directly with the ICE. Extracts were prepared from insulin-producing βTC-1 and HIT cells and non-insulin-producing HeLa and rat liver cells. Binding reactions were conducted in the presence of an extract, a ^{32}P-labeled probe corresponding to rat insulin II gene sequences from −104 to −85, and KCl at 50, 125, or 200 mM. Four protein-DNA complexes, which were labeled A to D in order of decreasing mobility, were detected in these binding experiments (Fig. 2). Protein-DNA complexes A and B were found in all of the cell extracts. All four protein-DNA complexes were specific, since their formation was prevented by cold probe sequences but not by nonspecific competitors (data not shown; Fig. 3). The B complex was detected only when binding reactions were conducted in 50 mM KCl, while formation of complex A was not nearly as salt dependent. Formation of protein-DNA complexes C and D was also dependent on the salt concentration in the binding reaction. In contrast to complex B formation, the C and D complexes were detected only at high salt concentrations (>100 mM). Figure 2B shows a longer exposure of the gel depicted in Fig. 2A. This exposure more clearly shows the presence of complex B and lack of complexes C and D in the HeLa extract. This exposure also shows the presence of complex D in the rat liver extract. Most interestingly, both C and D protein-DNA complexes are cell type specific, as they are uniquely formed in pancreatic β- and liver cell extracts,



Construct Transfected	Cell Line	Densitometric Intensity of Protected Fragment	Normalised Level
pSV -90	βTC	1.23	1.52
pSV -100	βTC	5.33	6.83
pSV -90	HeLa	3.13	4.12
pSV -100	HeLa	1.1	1.62

FIG. 1. Insulin gene −100 to −91 sequences activate transcription in a cell-specific manner when linked to a heterologous enhancer. (A) Structures of recombinant expression plasmids. pSV−90 consists of an SV40 enhancer-driven expression vector containing the coding sequences of the gene for CAT fused to the −90- to +8-base-pair sequence of the rat insulin II gene. SV40 2×72 enhancer sequences were subcloned from the pβ2×72$^-$ expression plasmid described by Ondek et al. (34). The shaded boxes represent SV40 enhancer sequences, and the thick solid lines refer to insulin sequences. Recombinant pSV−100 was constructed by fusing an oligonucleotide containing rat insulin II gene sequences from −100 to −91 (sequence shown above) to the 5′ end of the −90 rat insulin II gene promoter sequence. The triangle 3′ to the insulin gene transcription start site in pSV−100 represents an additional 6 base pairs within the polylinker of pSV−100. (B) Autoradiogram of RNase protection analyses. βTC-1 and HeLa cells were transfected with 5 μg of plasmid pSV−90 or pSV−100; an internal control plasmid, OVEC-REF (0.25 μg); or pUC19 DNA (4.75 μg). mRNA levels produced from these expression plasmids were assayed by RNase protection analyses. Protected fragments 286, 280, and 151 nucleotides long are expected for correctly initiated pSV−100, pSV−90, and OVEC-REF transcripts, respectively. The larger protected fragments from pSV−100, compared with pSV−90, are a result of the additional 6 base pairs within the polylinker of pSV−100. Lanes: 1, pSV−90 in βTC-1 cells; 2, pSV−100 in βTC-1 cells; 3, pSV−90 in HeLa cells; 4, pSV−100 in HeLa cells. The insulin signal shown is from a 1-day exposure of the RNase protection assay, while the OVEC-REF signal shown is from a 3-day exposure. (C) Quantitation of amounts of correctly initiated insulin transcripts. Relative amounts of correctly initiated insulin RNAs were determined by densitometric scanning of the autoradiogram and are expressed relative to the activity of the OVEC-REF internal control.

TABLE 1. Effects of point mutations within the −100 to −91 element on expression in β-cells[a]

Mutant nucleotide	Sequence from −100 to −91	Relative CAT enzyme activity	
		βTC-1 cells	HIT cells
Wild type	5′-GCCATCTGCT-3′	5.2	3.5
−100	5′-tCCATCTGCT-3′	5.5	2.8
−99	5′-GaCATCTGCT-3′	2.5	2.0
−98	5′-GCaATCTGCT-3′	2.5	2.0
−97	5′-GCCcTCTGCT-3′	1.5	0.8
−96	5′-GCCAgCTGCT-3′	4.8	2.8
−95	5′-GCCATaTGCT-3′	2.5	2.1
−94	5′-GCCATCgGCT-3′	1.0	0.9
−93	5′-GCCATCTtCT-3′	1.3	0.8
−92	5′-GCCATCTGaT-3′	1.3	0.8
−91	5′-GCCATCTGCg-3′	3.5	3.7
pSV−90	Not present	1.0	1.0

[a] The −100 to −91 element point mutations were constructed by fusing the mutant oligonucleotide to the end of the −90 rat insulin II gene promoter sequence as outlined in Fig. 1A. Values are presented relative to pSV−90 expression. The experiment was repeated on four separate occasions. The results shown are from one representative experiment. Standard deviations were less than 20%.

respectively. Thus, both cell-specific and ubiquitous factors specifically interact with the ICE. Furthermore, our results indicate that binding of cellular proteins to the ICE is salt dependent.

Previously, two other laboratories demonstrated that the rat insulin I gene cognate ICE bound to cellular proteins (30, 33). These studies did not make it clear whether the factors which bound to this element were unique to β-cells, as demonstrated by Ohlsson and Edlund (33), or ubiquitously distributed, as shown by Moss et al. (30). Our results perhaps explain the discrepancy between these studies. The experiments conducted by Moss et al. were performed at a low salt concentration (75 mM KCl), while Ohlsson and Edlund used a high salt concentration (150 mM KCl) in their binding assays. As we have shown here, use of a low salt concentration precludes detection of β-cell-type-specific protein-DNA complexes. It is surprising that Ohlsson and Edlund were unable to detect the A complex, which was detected in all of the extracts we assayed at a high salt concentration. This discrepancy could be due to differences in extract preparation procedures or, more probably, to the fact that slightly different DNA probes were used in the two studies.

Positive activation of the ICE appears to require β-cell-specific binding factors. Since a particular protein(s) capable of forming a protein-DNA complex (complex C) appeared to be uniquely present (active) in insulin-producing cells, we tested its ability to interact with wild-type and mutant forms of the ICE by using the gel mobility shift assay described above. If protein-DNA complex C is formed by the putative β-cell-specific IAF, it should not bind efficiently to mutant forms of the ICE which cannot activate transcription of the insulin gene in vivo (Table 1).

Binding experiments were performed with extracts prepared from βTC-1 cells (and HIT, HeLa and rat liver extracts; data not shown) and the ^{32}P-labeled wild-type ICE.

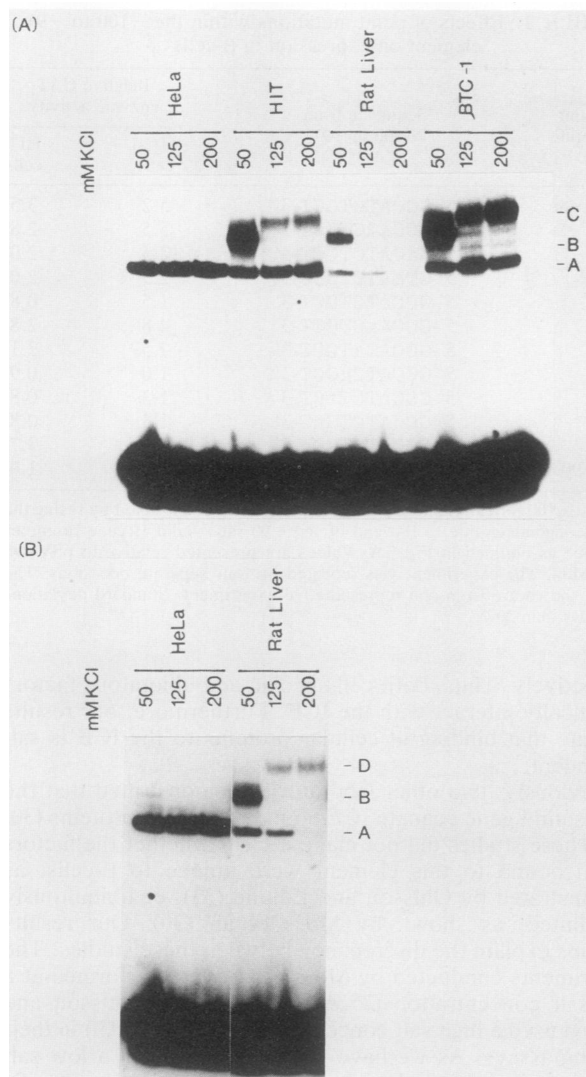


FIG. 2. Multiple cellular factors bind to the ICE. (A) Binding and gel electrophoresis conditions were as described by Carthew et al. (4). The final KCl concentration in the assay is shown for each lane. Nuclear extracts from HIT, βTC-1, and HeLa cells and rat liver were used. The major bands discussed in the text are labeled. The gel was exposed to autoradiography for 18 h. (B) A 96-h exposure of the HeLa cell and rat liver lanes.

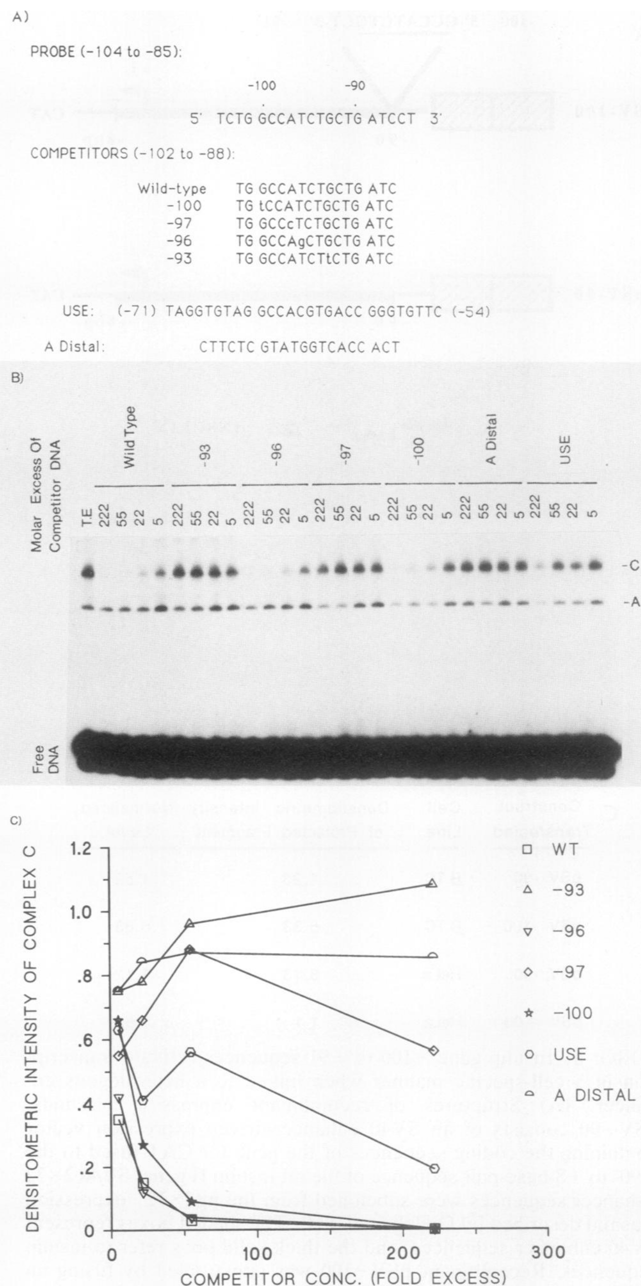


FIG. 3. Formation of complex C is specifically competed for by ICE mutants that mediate positive β-cell-type-specific activity. (A) Sequences of probe and competitor oligonucleotides. (B) Binding reactions were conducted in 200 mM KCl with whole βTC-1 cell extract. The molar ratio of competitor to probe is shown for each lane. The major bands are labeled as described in the legend to Fig. 2. (C) Quantitation of the effect of ICE mutants on complex C formation. The relative amount of complex C in each lane was determined by densitometric scanning of the autoradiogram. WT, Wild type; conc., concentration.

These binding reactions were performed in 50 or 200 mM KCl with various cold competitor DNAs at molar ratios of 5, 22, 55, and 222 to 1 (competitor/probe ratios). The sequences of the probe and competitor DNAs used for this experiment are shown in Fig. 3A. The experimental results presented in Fig. 3B and C were obtained with 200 mM KCl. Under these conditions, complexes A and C were the most abundant (Fig. 2). The specificity of these interactions was shown by the fact that a specific (i.e., wild type) but not a nonspecific (i.e., A Distal) competitor could abolish formation of complex C. Oligonucleotides bearing the sequence of the adenovirus type 2 major late promoter upstream element (USE) competed for formation of complex C, albeit inefficiently. This observation is examined in more detail below. Moreover, the patterns of competition with the various alleles of the ICE were exactly those predicted. Thus, mutant forms of the ICE which fail to activate insulin gene transcription in vivo are not efficient competitors for formation of complex C in vitro. Neither the −93 nor the −97 mutant form of the ICE competed well, while both the −96 and −100 mutant ICEs did compete efficiently against wild-type ICE sequences (Fig. 3B and C). Analogous results were obtained with complex C formed in HIT cell extracts (data not shown). None of the other protein-DNA complexes (in βTC-1, HIT,

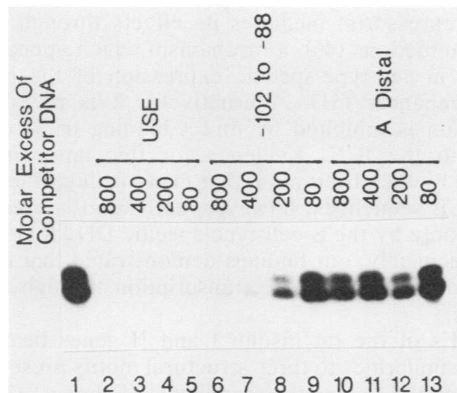


FIG. 4. Both USE and ICE sequences specifically inhibit transcription of the adenovirus type 2 major late promoter in vitro. Transcription reactions were conducted as previously described (49). The sequences of the USE, ICE, and A Distal competitor oligonucleotides are shown in Fig. 3A. The molar excess of competitor DNA present in the preincubation of extract and template DNAs is shown above each lane. Specific transcription initiation from the adenovirus 2 type major late promoter plasmid (pSmaF) was measured by primer extension. Adenovirus type 2 major late promoter transcription was monitored with a primer (5'-AGGGAG TACTCACCCCAACAGCTGGCCCTC-3') complementary to bases +25 to +54 of the noncoding strand of the major late promoter. An extended product of 54 nucleotides is expected for adenovirus major late promoter-specific transcripts.

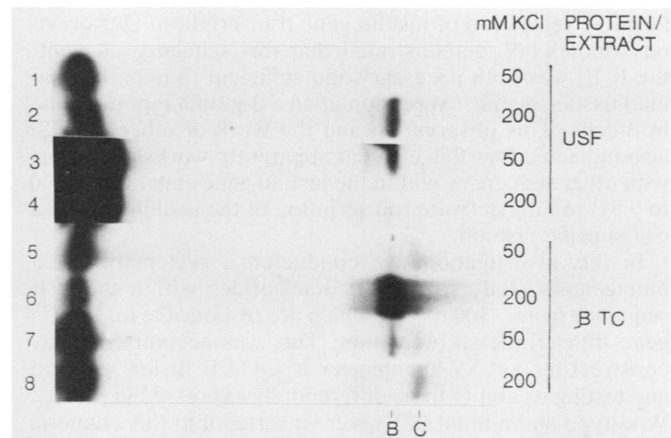


FIG. 5. The in vitro-translated USF binds to the ICE. Binding reactions were conducted as described in the legend to Fig. 2. The USE (lanes 1, 2, 5, and 6) or −104 to −85 (lanes 3, 4, 7, and 8) labeled probe was incubated with either the reticulocyte lysate-translated human USF or βTC-1 extract in 50 or 200 mM KCl as indicated. The bands are labeled as in Fig. 2. Lanes 1, 2, and 5 to 8 are from a 1-day exposure, while lanes 3 and 4 are from a 5-day exposure.

HeLa, or rat liver extracts) formed under these (i.e., 200 mM KCl) or other conditions (lower ionic strengths) exhibited the competition patterns described above for complex C. Together, these data strongly suggest that the protein which forms protein-DNA complex C is a β-cell-specific IAF whose activity is manifest through ICE sequences −100 to −91.

The USF binds to the ICE. The data from the preceding experiments indicated that both cell-type-specific and non-cell-type-specific factors bound to the ICE. It appears that binding of a β-cell-specific factor(s) to the ICE is important for cell-type-specific expression; however, the functions of the other proteins which interact with this element are unknown. As shown in the gel mobility shift assay presented in Fig. 3, DNA containing the binding site for USF competed with the −104 to −85 wild-type probe for the A and C protein-DNA complexes. To obtain a level of competition comparable to that of the wild-type probe, a 10-fold increase in the molar amount of USE sequences was used (data not shown). These results indicated that the USF might interact with ICE sequences in vitro. This result is somewhat surprising, since the core sequences of the USE and the ICE share only limited sequence homology.

Because of the apparent cross talk (specific or nonspecific) between the USE and ICE activator factors, we examined whether an oligonucleotide containing ICE sequences would affect transcription of the adenovirus type 2 major late transcription promoter in vitro. Thus, if the USF bound to the ICE, then the presence of an ICE oligonucleotide in the assay would sequester and, as a consequence, reduce adenovirus type 2 major late promoter transcription. The effects of adding increasing amounts of the USE, the ICE, and nonspecific, double-stranded oligonucleotides on adenovirus type 2 major late promoter transcription at 80, 200, 400, and 800-fold molar excesses were tested (Fig. 4). The USE competitor resulted in a marked decrease in the level of adenovirus type 2 major late promoter transcription at all of the competitor concentrations tested (Fig. 4, lanes 2 to 5). Similarly, the ICE oligonucleotide reduced transcription, but less efficiently, whereas a nonspecific DNA competitor did not (Fig. 4; compare lanes 6 to 9 with 10 to 13). These results further support the observation that the USF binds to the ICE.

To test directly whether the USF bound to the ICE, a clone encoding the human USF was expressed in vitro. The DNA-binding properties of this USF were examined in the gel mobility shift assay. Binding reactions were performed in either 50 or 200 mM KCl with similarly sized USE and ICE probes. The results of this experiment are presented in Fig. 5. The USF bound to both probes to yield the B complex described for the −104 to −85 probe in Fig. 2. Interestingly, binding of the USF to these probes was dependent on the KCl concentrations in the binding reactions. The authentic USF bound most efficiently to USE sequences at high salt concentrations, whereas it bound the ICE probe only at low salt concentrations (the conditions of the in vitro assay). An identical gel mobility shift pattern for the B complex was found with βTC-1 extracts (Fig. 5). Since similar amounts of in vitro-synthesized protein were used in the experiments described in Fig. 5, the intensity of the B band indicates that USF binds with higher affinity to the USE than to the ICE. These results indicate that the USF binds in vitro to the ICE. The importance of this observation to insulin gene regulation is unclear. We do not know whether in vivo binding of the USF has any role in ICE regulation. The USF cannot impart β-cell-type activation, since changes in the ICE sequences to make it more USE-like eliminated cell-type-specific expression in vivo (data not shown). However, one could speculate that the USF plays a role in negative control of ICE expression (see Discussion).

DISCUSSION

The studies described in this report were undertaken to characterize in detail the rat insulin II gene regulatory sequences which map between nucleotides −100 and −91. This sequence element is the target for both positive and

negative regulation of insulin gene transcription. Our previous studies (49) demonstrated that this sequence element, the ICE, was both necessary and sufficient to impart (minimal) tissue-specific expression upon a β-globin reporter gene in β-cells. This observation and the work of others (9, 22) also indicated that this element apparently works in concert with other sequences within the insulin gene enhancer (−340 to −91) to fully activate transcription of the insulin gene in a cell-specific fashion.

In this investigation, we conducted a systematic point mutagenesis study to identify nucleotides within the ICE sequence from −100 to −91 which are responsible for insulin gene transcriptional activation. This was accomplished by constructing an SV40-enhancer-ICE-CAT fusion plasmid and testing its ability to be differentially expressed in β-cells. Wild-type and mutant ICEs were inserted into this chimeric CAT expression plasmid. Transfection experiments indicated that unique base pairs within the ICE were involved in β-cell-specific insulin gene transcriptional activation. Nucleotides −97, −94, −93, and −92 were found by this in vivo functional test to be crucial for β-cell-specific expression.

This information provided us the biological framework with which to begin in vitro studies to attempt to identify proteins which specifically bound to the ICE. Proteins which specifically bound to the ICE would be candidates for insulin gene *trans*-acting transcriptional activators. The gel mobility shift assay was used for this purpose. Our initial experiments indicated that multiple protein-DNA complexes were formed with the ICE in vitro. However, the pattern of complexes formed with this DNA fragment was highly variable. Detailed optimization experiments indicated that complex formation was very salt dependent.

Under high-ionic-strength conditions (200 mM KCl), we observed a protein-DNA complex (termed complex C) which was formed only by proteins present in extracts prepared from insulin-producing cells. This result suggested that the protein forming this complex could be an insulin gene-specific *trans*-activator. This hypothesis was tested by examining the ability of this β-cell-specific protein to interact with the family of wild-type and mutant ICEs used for our in vivo studies. If the β-cell C-complex-forming protein were the insulin gene-specific *trans*-activator, it should interact differentially with wild-type and mutant ICEs. Specifically, ICE down mutations (i.e., −97, −93 ICE mutant alleles) should decrease binding, while ICE silent mutations (i.e., −100 and −96 mutant ICE alleles) should not affect the affinity of binding of this protein to the ICE. This was exactly the pattern of protein-DNA binding observed (Fig. 3). These results strongly suggest that the protein(s) responsible for forming protein-DNA complex C is in fact a β-cell-specific factor which is involved in activating insulin gene transcription through ICE sequences −100 to −91. This factor has been termed the IAF. We are currently attempting to purify and characterize the IAF by direct biochemical methods (see also below).

Control of cell-type-specific gene expression through distinct *cis*-acting sequences and *trans*-acting factors has been described in a number of systems. For the insulin gene enhancer system itself, it has been clearly shown that ICE sequences are critical for β-cell-specific expression. The existence of a uniquely active-distributed β-cell-type-specific positive regulator was demonstrated in this study; however, we have previously proposed that negative regulators of ICE expression are present in non-β-cells.

The mechanism(s) involved in negative control is unclear (for a recent review, see reference 25). We have suggested that the repressor(s) mediates its effects through protein-protein interactions (46), a mechanism which appears to be important in cell-type-specific expression of the immunoglobulin enhancer (51). Alternatively, it is possible that transcription is inhibited by direct binding of a repressor factor(s) to the ICE. Evidence for this latter model is supported by the observation that many cellular factors can bind to ICE sequences; however, activation is presumably mediated only by the β-cell-type-specific DNA-binding factor(s). Specifically, our findings demonstrated that the USF could bind but not activate transcription through the ICE sequences.

The ICEs of the rat insulin I and II genes bear strong sequence similarities to three structural motifs present in the immunoglobulin heavy-chain enhancer (E boxes) (30). Mutation of the ICE to an immunoglobulin heavy-chain μE2 enhancer element allows it to function like the insulin element in vivo (30). (The sequence of the immunoglobulin heavy-chain μE2 enhancer element is identical to that of the −100 to −91 element, except for a single-base-pair mutation at nucleotide −96.) These results led Moss et al. (30) to propose that the immunoglobulin heavy-chain μE2 enhancer element-binding factors may be involved in insulin gene expression. Since the results presented here suggest that the activator factor(s) is uniquely present in β-cells, we feel that this hypothesis is incorrect. We instead propose that the mutation in the ICE at base pair −96 simply does not alter the interaction of the β-cell-type-specific IAF with this element (Fig. 3B and C). However, these results may indicate the DNA-binding domains of these various factors (i.e., the immunoglobulin heavy-chain μE2 enhancer element factor, USF, and IAF) are related.

Very recently, evidence has been presented by Murre et al. (31) demonstrating that proteins that bind to the κ-μE2 box element of the immunoglobulin enhancer share a common amphipathic helix-loop-helix (HLH) structural motif. This conserved HLH region is found in a wide variety of cellular proteins, including daughterless, achaetescute T3, and MyoD, as well as proteins believed to be involved in immunoglobulin activation, E12 and E47 (31, 32). The HLH motif is essential for dimerization and DNA binding (32). We are currently screening a human λgt11 cDNA library derived from a human β-cell insulinoma for the IAF. Since other proteins that bind to the ICE element contain the HLH motif, one might anticipate finding such a motif within the IAF. Since proteins containing the HLH structural motif can form DNA-binding heterodimers (32), one could envision that negative regulation of transcription mediated by the ICE could occur if the activator factor formed a heterodimer with a protein(s) which, upon binding, inhibited insulin gene transcription.

These data indicate the complexity of regulatory control mediated by the ICE. An understanding of the regulatory mechanism(s) that mediates cell type-specific expression of this element in particular and the insulin gene in general will require isolation and characterization of both the negative- and positive-acting cellular regulators that interact with these DNA sequences. Such studies are currently in progress in our laboratories.

ACKNOWLEDGMENTS

We thank Teresa Lavenue and Ian Jones for excellent technical assistance. We are greatly indebted to Bob Roeder and Polly Gregor for providing us with the human USF expression vector.

J.W. was supported by a postdoctoral fellowship from the Juvenile Diabetes Foundation. R.S. is a recipient of a career develop-

ment award from the Juvenile Diabetes Foundation. This work was supported by grant 188773 from the Juvenile Diabetes Foundation and Public Health Service grants GM-30257 (to R.S.) and GM-40517 (to P.A.W.), both from the National Institutes of Health. Partial support was also derived from the Vanderbilt University Diabetes Research and Training Center Molecular Biology Core Laboratory (Public Health Service grant P60 DK20593 from the National Institutes of Health).

LITERATURE CITED

1. **Angel, P., M. Imagawa, R. Chiu, B. Stein, R. J. Imbra, H. J. Rahmsdorf, C. Jonat, P. Herrlich, and M. Karin.** 1987. Phorbol ester-inducible genes contain a common cis element recognized by a TPA-modulated trans-acting factor. Cell **49**:729–739.
2. **Barberis, A., S. F. Giulio, and M. Busslinger.** 1987. Mutually exclusive interaction of the CCAAT-binding factor and of a displacement protein with overlapping sequences of a histone gene promoter. Cell **50**:347–359.
3. **Briggs, M. R., J. T. Kadonaga, S. P. Bell, and R. Tjian.** 1986. Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. Science **234**:47–52.
4. **Carthew, R. W., L. A. Chodosh, and P. A. Sharp.** 1985. An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major late promoter. Cell **43**:439–448.
5. **Chan, S. J., V. Episkopou, S. Zeitlin, S. K. Karathanasis, A. MacKrell, D. F. Steiner, and A. Efstratiadis.** 1984. Guinea pig preproinsulin gene: an evolutionary compromise? Proc. Natl. Acad. Sci. USA **81**:5046–5050.
6. **Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18**:5294–5299.
7. **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.
8. **Cordell, B., D. Diamond, S. Smith, J. Punter, H. H. Schone, and H. M. Goodman.** 1982. Disproportionate expression of the two insulin nonallelic rat insulin genes in a pancreatic tumor is due to translational control. Cell **31**:531–542.
9. **Crowe, D. T., and M. Tsai.** 1989. Mutagenesis of the rat insulin II 5'-flanking region defines sequences important for expression in HIT cells. Mol. Cell. Biol. **9**:1784–1789.
10. **Edlund, T., M. D. Walker, P. J. Barr, and W. J. Rutter.** 1985. Cell-specific expression of the rat insulin gene: evidence for the role of two distinct 5' flanking sequences. Science **230**:912–916.
11. **Efrat, S., S. Linde, H. Kofod, D. Spector, M. Delannoy, S. Grant, D. Hanahan, and S. Baekkeskov.** 1988. Beta cell lines derived from transgenic mice expressing a hybrid insulin gene-oncogene. Proc. Natl. Acad. Sci. USA **85**:9037–9041.
12. **Ephrussi, A., G. M. Church, S. Tonegawa, and W. Gilbert.** 1985. B-lineage specific interactions of an immunoglobulin enhancer with cellular factors in vivo. Science **227**:134–140.
13. **Episkopou, V., A. J. M. Murphy, and A. Efstratiadis.** 1984. Cell-specific expression of a selectable hybrid gene. Proc. Natl. Acad. Sci. USA **81**:4657–4661.
14. **Goodbourn, S., H. Burstein, and T. Maniatis.** 1986. The human β-interferon gene enhancer is under negative control. Cell **45**:601–610.
15. **Gorski, K., M. Carneiro, and U. Schibler.** 1986. Tissue-specific in vitro transcription from the mouse albumin promoter. Cell **47**:767–776.
16. **Hanahan, D.** 1985. Heritable formation of pancreatic β-cell tumours in transgenic mice expressing recombinant insulin/simian virus 40 oncogenes. Nature (London) **315**:115–122.
17. **Hatzopoulos, A. K., U. Schlokat, and P. Gruss.** 1988. Enhancers and cis-acting regulatory sequences, p. 43–96. *In* B. D. Hames and D. M. Glover (ed.), Eucaryotic RNA synthesis and processing, horizons in molecular biology series. IRL Press, Oxford.
18. **Hoeffler, W. K., R. Kovelman, and R. G. Roeder.** 1988. Activation of transcription factor IIIC by the adenovirus E1A protein. Cell **53**:907–920.
19. **Johnson, P. F., W. H. Landschulz, B. J. Graves, and S. L. McKnight.** 1987. Identification of a rat liver protein that binds to the enhancer core element of three animal viruses. Genes Dev. **1**:133–146.
20. **Jones, K. A., J. T. Kadonaga, P. J. Rosenfeld, T. J. Kelly, and R. Tjian.** 1987. A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. Cell **48**:79–89.
21. **Karlsson, O., T. Edlund, J. B. Moss, W. J. Rutter, and M. D. Walker.** 1987. A mutational analysis of the insulin gene transcription control region: expression in beta cells is dependent on two related sequences within the enhancer. Proc. Natl. Acad. Sci. USA **84**:8819–8823.
22. **Karlsson, O., M. D. Walker, W. J. Rutter, and T. Edlund.** 1989. Individual protein-binding domains of the insulin gene enhancer positively activate β-cell-specific transcription. Mol. Cell Biol. **9**:823–827.
23. **Landolfi, N. F., J. D. Capra, and P. W. Tucker.** 1986. Interaction of cell-type-specific nuclear proteins with immunoglobulin V_H promoter region sequence. Nature (London) **323**:548–551.
24. **Lee, W., P. Mitchell, and R. Tjian.** 1987. Purified transcription factor AP-1 interacts with TPA-inducible elements. Cell **49**:741–751.
25. **Levine, M., and J. L. Manley.** 1989. Transcriptional repression of eukaryotic promoters. Cell **59**:405–408.
26. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
27. **Maniatis, T., S. Goodbourne, and J. A. Fischer.** 1987. Regulation of inducible and tissue-specific gene expression. Science **236**:1237–1244.
28. **McKnight, S. L., E. R. Gavis, and R. Kingsbury.** 1981. Analysis of transcription regulatory signals of the HSV thymidine kinase gene: identification of an upstream control region. Cell **25**:385–398.
29. **Melton, D. A., P. H. Kreig, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12**:7035–7056.
30. **Moss, L. G., J. B. Moss, and W. J. Rutter.** 1988. Systematic binding analysis of the insulin gene transcription control region: insulin and immunoglobulin enhancers utilize similar transactivators. Mol. Cell. Biol. **8**:2620–2627.
31. **Murre, C., P. S. McCaw, and D. Baltimore.** 1989. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. Cell **56**:777–783.
32. **Murre, C., P. S. McCaw, H. Vaessin, M. Cudy, L. Y. Jan, Y. N. Jan, C. V. Cabrera, J. Buskin, S. D. Hauschka, A. B. Lassar, H. Weintraub, and D. Baltimore.** 1989. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. Cell **58**:537–544.
33. **Ohlsson, H., and T. Edlund.** 1988. A β-cell-specific protein binds to two major regulatory sequences of the insulin gene enhancer. Proc. Natl. Acad. Sci. USA **85**:4228–4231.
34. **Ondek, B., A. Shepard, and W. Herr.** 1987. Discrete elements within the SV40 enhancer region display different cell-specific enhancer activities. EMBO J. **6**:1017–1025.
35. **Pettersson, M., and W. Schaffner.** 1987. A purine-rich DNA sequence motif present in SV40 and lymphotropic papovavirus binds a lymphoid-specific factor and contributes to enhancer activity in lymphoid cells. Genes Dev. **1**:962–972.
36. **Plumb, M. A., V. V. Lobanenkov, R. H. Nicolas, C. A. Wright, S. Zavon, and G. H. Goodwin.** 1986. Characterization of chicken erythroid proteins which bind to the nuclease hypersensitive regions upstream of the βA- and βH-globin genes. Nucleic Acids Res. **19**:7675–7693.
37. **Prywes, R., and R. G. Roeder.** 1986. Inducible binding of a factor to the c-fos enhancer. Cell **47**:777–784.
38. **Sakonju, S., and D. D. Brown.** 1982. Contact points between a positive transcription factor and the Xenopus 5S RNA gene. Cell **31**:395–405.
39. **Sawadogo, M., and R. G. Roeder.** 1985. Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. Cell **43**:165–175.

40. **Scholer, H., and P. Gruss.** 1985. Cell-type specific enhancement in vitro requires the presence of trans-acting factors. EMBO J. **4**:3005–3013.
41. **Sen, R., and D. Baltimore.** 1986. Multiple nuclear factors interact with the immunoglobulin enhancer sequences. Cell **46**:705–716.
42. **Sen, R., and D. Baltimore.** 1986. Inducibility of κ immunoglobulin enhancer-binding protein NF-κB by a posttranslational mechanism. Cell **47**:921–928.
43. **Shapiro, D. J., P. A. Sharp, W. W. Wahli, and M. Keller.** 1988. A high-efficiency HeLa cell nuclear transcription extract. DNA **1**:47–55.
44. **Sorger, P. K., M. J. Lewis, and H. R. B. Pelham.** 1987. Heat shock factors regulated differentially in yeast and HeLa cells. Nature (London) **329**:81–84.
45. **Staudt, L. M., H. S. 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.
46. **Stein, R. W., and J. Whelan.** 1989. Insulin gene activity is inhibited by adenovirus 5 E1a gene products. Mol. Cell. Biol. **9**:4531–4534.
47. **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.
48. **Westin, G., T. Gerster, M. M. Muller, G. Schaffner, and W. Schaffner.** 1987. OVEC, a versatile system to study transcription in mammalian cells and cell-free extracts. Nucleic Acids Res. **15**:6787–6798.
49. **Whelan, J., D. Poon, P. A. Weil, and R. Stein.** 1989. Pancreatic β-cell-type-specific expression of the rat insulin II gene is controlled by positive and negative transcriptional elements. Mol. Cell. Biol. **9**:3253–3259.
50. **Xiao, J. H., I. Davidson, D. Ferrandon, R. Rosales, M. Vigneron, M. Macchi, F. Ruffenach, and P. Chambon.** 1987. One cell specific and three ubiquitous nuclear proteins bind in vitro to overlapping motifs in domain B1 of the SV40 enhancer. EMBO J. **6**:3015–3025.
51. **Yu, H., B. Porton, L. Shen, and L. A. Eckhardt.** 1989. Role of the octamer motif in hybrid cell extinction of immunoglobulin gene expression: extinction is dominant in a two enhancer system. Cell **58**:441–448.
52. **Zinn, K., D. Di Maio, and T. Maniatis.** 1983. Identification of two distinct regulatory regions adjacent to the human β-interferon gene. Cell **34**:865–879.