MOLECULAR AND CELLULAR BIOLOGY, Apr. 1990, p. 1784–1788
0270-7306/90/041784-05$02.00/0

Vol. 10, No. 4

Cooperativity of Sequence Elements Mediates Tissue Specificity of the Rat Insulin II Gene

YOUNG-PING HWUNG, YI-ZHONG GU, AND MING-JER TSAI*

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

Received 20 September 1989/Accepted 14 December 1989

The 5′-flanking region of the rat insulin II gene (−448 to +50) is sufficient for tissue-specific expression. To further determine the tissue-specific *cis*-acting element(s), important sequences defined by linker-scanning mutagenesis were placed upstream of a heterologous promoter and transfected into insulin-producing and -nonproducing cells. Rat insulin promoter element 3 (RIPE3), which spans from −125 to −86, was shown to confer β-cell-specific expression in either orientation. However, two subregions of RIPE3, RIPE3a and RIPE3b (defined by linker-scanning mutations), displayed only marginal activities. These results suggest that the two subregions cooperate to confer tissue specificity, presumably via their cognate binding factors.

The insulin gene is expressed exclusively in pancreatic β cells in normal adults and is therefore a good model system for the study of tissue-specific gene expression. Previous studies have shown that less than 600 base pairs (bp) of the 5′-flanking sequences are sufficient for tissue specificity of the insulin gene in both cultured cells (6, 23) and transgenic mice (10). Therefore, the interaction between upstream sequence elements and *trans*-acting factors is important for regulation of insulin gene expression. Upstream flanking sequences of the mammalian insulin genes are homologous up to 500 bp (22). Furthermore, interspecies gene transfer experiments indicate that the molecular mechanism underlying the tissue specificity of the insulin gene is conserved among mammals.

There are two nonallelic insulin genes in rats, the rat insulin I (rInsI) and rat insulin II (rInsII) genes (3, 23). These two genes have more than 80% sequence similarity in both the coding and flanking sequences, and they are coordinately expressed (8). The rInsII gene is more similar in structure than the rInsI gene to the other mammalian insulin genes, and it is proposed that rInsI gene was duplicated from the ancestral rInsII gene (21). To better define which sequences within the 5′-flanking region of the rInsII gene are important for gene activity, linker-scanning (LS) mutagenesis was carried out in our laboratory (4). Mutation of seven regions (referred to as the rat insulin promoter elements [RIPEs]) resulted in significant decreases in gene activity, suggesting that the insulin gene control region is composed of several functional domains or *cis* elements. These *cis* elements may cooperate through their cognate *trans*-acting factors to achieve high levels of transcription. One of the *trans*-acting factor is the COUP (chicken ovalbumin upstream promoter) transcription factor (18, 19), which binds to RIPE1 (−53 to −46) and is likely to play an important role in insulin promoter function (11). Interestingly, the COUP transcription factor contacts the insulin and ovalbumin promoter quite differently (12), the physiological significance of which remains unknown.

These mutagenesis results, however, did not define the control element(s) for tissue specificity. It is possible that one or a combination of the LS-defined sequence elements can function only in insulin-producing cells, thus conferring tissue specificity. To investigate this possibility, we inserted the important sequence elements in front of a heterologous promoter and examined expression levels in insulin-producing and -nonproducing cell types.

The two most critical regions in the 5′-flanking sequences for rInsII gene expression are −101 to −92 and −124 to −111 (RIPE3a and RIPE3b, respectively), as shown by our LS mutagenesis. Replacement of either region with a *Sac*I linker yielded only 4% of the wild-type gene activity (4). We therefore tested the tissue specificities of these sequences. The vector used in this assay, pOVCAT-50, contains the chicken ovalbumin gene sequence from −50 to +43, in which there is no known sequence element other than the TATA box. The ovalbumin minimal promoter is fused to a reporter gene that contains the coding region of the chloramphenicol acetyltransferase (CAT) gene and simian virus 40 splice-polyadenylation sequences (Fig. 1A). Three oligonucleotides, covering RIPE3a (−110 to −86), RIPE3b (−125 to −101), and a mutated sequence of RIPE3b (mRIPE3b; −156 to −101, in which the sequence from −125 to −111 was replaced by a *Sac*I linker), respectively, were inserted in various copy numbers and orientations into the *Bgl*II site of pOVCAT-50. The resulting plasmids were transiently transfected into HIT-T15 M2.2.2 cells, a hamster insulin tumor (HIT) cell line (20), or BHK cells, a fibroblast cell line.

HIT-T15 M2.2.2 or BHK cells were cultured at 37°C in Dulbecco modified Eagle medium supplemented with 10% fetal bovine serum, 100 U of penicillin per ml, and 100 μg of streptomycin per ml in the presence of 5% CO_2. Transient transfections (5) and CAT assays (9) were performed as previously described.

The results (Fig. 2 and Table 1) indicate that RIPE3a and RIPE3b oligonucleotides were both active in HIT cells. In contrast, an oligonucleotide carrying a mutated RIPE3b (mRIPEb) sequence, in which the natural sequences from −125 to −111 were replaced by a *Sac*I linker, resulted in no activation. Furthermore, RIPE3b functioned in both orientations, whereas RIPE3a was active only in the natural orientation. Neither oligonucleotide stimulated pOVCAT-50 activity in BHK cells, suggesting that RIPE3a and RIPE3b are tissue-specific *cis* elements. Nevertheless, the activity of RIPE3a or RIPE3b was much less than that of the wild-type rInSII plasmid, pINSCAT448$^-$, which has 448 bp of rInsII gene 5′-flanking sequences (Fig. 2, lane 1; Table 1), indicating that neither sequence alone is sufficient for full activity of the gene.

* Corresponding author.

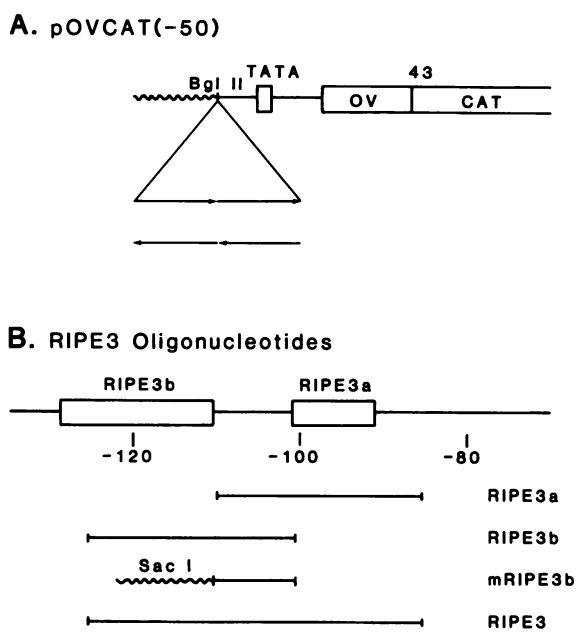

FIG. 1. Structures of pOVCAT-50 and oligonucleotides used. (A) Vector plasmid pOVCAT-50, which has the minimal promoter (−50 to +43) of the chicken ovalbumin gene (OV) linked to the coding region of the bacterial CAT gene and simian virus 40 splice-polyadenylation signal. The wavy line represents pBR322 sequences. Oligonucleotides were inserted into the *Bgl*II site in various copy numbers and orientations. (B) Extents of oligonucleotides (RIPE3a, RIPE3b, and RIPE3) used.

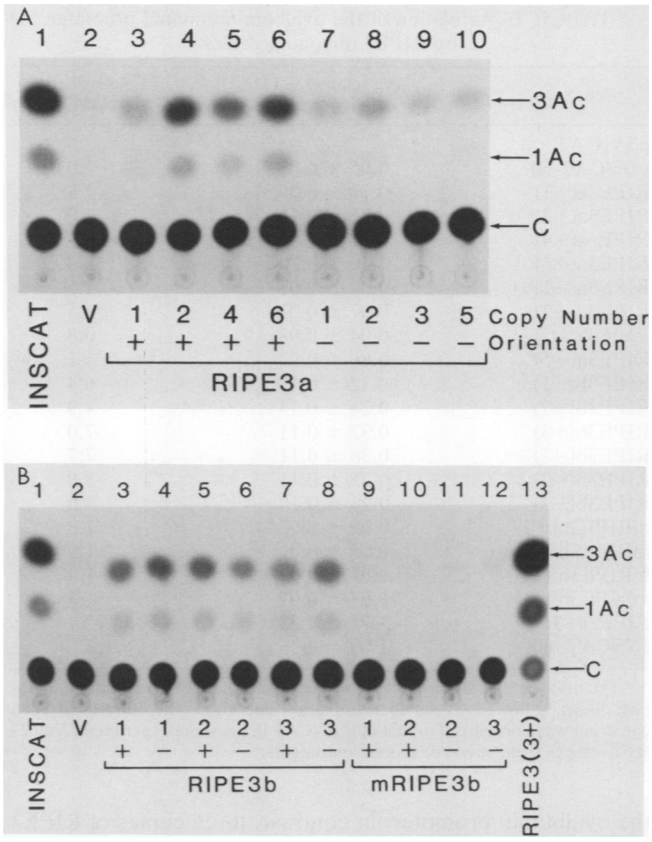

FIG. 2. Activity of RIPE3a and RIPE3b in HIT cells. Symbols and abbreviations: +, natural orientation; −, reverse orientation; 3-AC and 1-AC, 3- and 1-monoacetylated chloramphenicol, respectively; C, chloramphenicol. Plasmids used: (A) pINSCAT448$^-$, the wild-type control containing 448 bp of rInSII 5′-flanking sequences (lane 1), the vector pOVCAT-50 (lane 2), and RIPE3a inserted into pOVCAT-50 in various orientations and copy numbers (lanes 3 to 10); (B) pINSCAT448$^-$ (lane 1), pOVCAT-50 (lane 2), RIPE3b inserted into pOVCAT-50 in various orientations and copy numbers (lanes 3 to 8), mutated RIPE3b inserted into pOVCAT-50 in various orientations and copy numbers (lanes 9 to 12), and RIPE3(3+) (lane 13).

Since RIPE3a and RIPE3b are located next to each other and neither one alone can bestow full gene activity, it is possible that RIPE3a and RIPE3b are functionally dependent on each other. To test this hypothesis, an oligonucleotide containing RIPE3 (−125 to −86; Fig. 1B), which covers both regions, was synthesized, cloned, and used in transfection assays as described above. One copy of RIPE3 oligonucleotide had little effect, whereas two copies increased CAT activity significantly; three copies of RIPE3, in either orientation, activated expression to the level of the wild-type plasmid pINSCAT448$^-$ (Fig. 3). These results suggest that RIPE3 can activate a heterologous promoter in HIT cells in the absence of any other upstream promoter element. The activation is dependent on the copy number but not the orientation of RIPE3.

It is not surprising that more than one copy of RIPE3 was required to achieve the activity of rInsII full-length enhancer and promoter, whereas there is only one copy in the native sequences. As has been shown for the simian virus 40 enhancer, multimerization of enhancer elements that appear only once in the wild-type enhancer reconstitutes an active enhancer (references 7 and 16 and references therein). Similarly, duplication of the steroid hormone-responsive elements can eliminate the requirement of other *cis*-acting elements in stimulating transcription from the TATA box (2). Our LS mutagenesis result indicates that at least seven sequence elements are important for rInSII expression. It is conceivable that RIPE3 interacts with the other elements in the native sequences. These other elements, except RIPE1, are not characterized yet, and they may be important for physiological regulation of insulin gene activity, such as glucose or cyclic AMP control.

To ascertain whether the increase in CAT activity induced by RIPE3 actually reflects an increase of correctly initiated transcripts, the transcription start site was determined by RNase mapping. Total RNA was harvested (1) and treated (4) as previously described. The probe used was an antisense RNA covering the ovalbumin gene sequences between −50 and +43, as well as 250 bp of the CAT sequences (Fig. 4A). The correctly initiated transcript, when hybridized to the probe, should protect a fragment of 293 nucleotides from RNase digestion. In each experiment, the same amount of pSV_2CAT was cotransfected with the test plasmid to serve as an internal control. The transcripts of pSV_2CAT should hybridize to the CAT, but not the ovalbumin, sequences, to generate a protected band of 250 nucleotides (Fig. 4A).

Yeast RNA did not give any protected product in this assay, indicating that the experimental conditions are specific to the CAT gene transcript (Fig. 4B, lane 2). The RNA harvested from cells transfected with pSV_2CAT alone yielded a single band at the expected position. This internal control band appeared at similar intensities in all experiments, indicating that transfection efficiencies were rather constant. RNA from pOVCAT-50-transfected cells yielded no protected band of the expected size for transcript from

TABLE 1. Activation of the ovalbumin minimal promoter by RIPE3 oligonucleotides

Construct	Avg relative CAT activity ± SD[a]	Fold induction over pOVCAT-50
pSV_2CAT	1.00	
pOVCAT-50	0.05 ± 0.03	1.0
RIPE3a(+1)	0.17 ± 0.09	3.4
RIPE3a(+2)	0.35 ± 0.15	7.0
RIPE3a(+4)	0.33 ± 0.23	6.7
RIPE3a(+6)	0.40 ± 0.26	8.0
RIPE3a(−1)	0.07 ± 0.03	1.5
RIPE3a(−2)	0.07 ± 0.03	1.5
RIPE3a(−3)	0.04 ± 0.03	0.8
RIPE3a(−5)	0.06 ± 0.02	1.1
RIPE3b(+1)	0.32 ± 0.17	6.4
RIPE3b(+2)	0.24 ± 0.13	4.9
RIPE3b(+3)	0.32 ± 0.11	7.0
RIPE3b(−1)	0.38 ± 0.11	7.7
RIPE3b(−2)	0.25 ± 0.11	5.0
RIPE3b(−3)	0.23 ± 0.10	4.6
mRIPE3b(+1)	0.09 ± 0.07	1.8
mRIPE3b(+2)	0.09 ± 0.07	1.9
mRIPE3b(−2)	0.07 ± 0.07	1.4
mRIPE3b(−3)	0.07 ± 0.07	1.5
RIPE3(+3)	3.71 ± 2.53	75
INSCAT 448−	1.09 ± 0.26	22

[a] Quantitated by cutting individual spots from the thin-layer chromatography plates, followed by liquid scintillation counting. The activity of each construct was normalized to that of pSV_2CAT in the same experiment. Values are averages from seven or more experiments.

the ovalbumin promoter; in contrast, three copies of RIPE3 induced a strong protected band of 293 nucleotides. pRIPE3(1+), which had only background CAT activity, yielded no correctly initiated transcript. In addition to the correct transcripts, the pOVCAT-50 derivatives generated other bands, which did not appear in the pSV_2CAT-transfected cells. These bands may have arisen from unknown initiation sites in pOVCAT-50, and some of these transcripts may have been translated into functional CAT enzymes, resulting in the low CAT activity seen in pOVCAT-50-transfected cells (Fig. 2). Nevertheless, these bands did not change in either size or intensity when RIPE3 was inserted into pOVCAT-50. Taken together, these results suggest that RIPE3 activates transcription from the native initiation site.

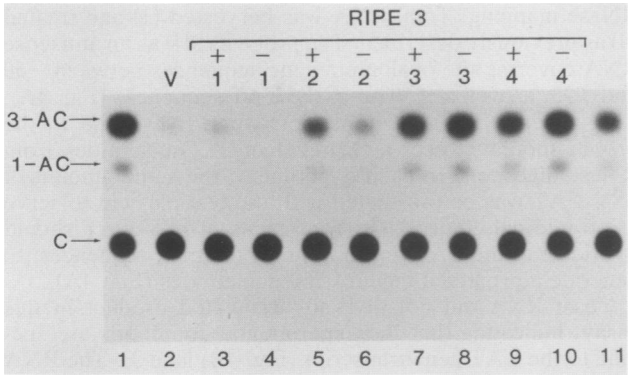

FIG. 3. Activity of RIPE3 in HIT cells. HIT cells were transfected with pSV_2CAT (lane 1), the vector (V) pOVCAT-50 (lane 2), pRIPE3(1+) through pRIPE3(4−) (lanes 3 through 10), and the wild-type control pINSCAT448$^-$, which has the full-length rInsII promoter and enhancer (lane 11). Symbols and abbreviations are as for Fig. 2.

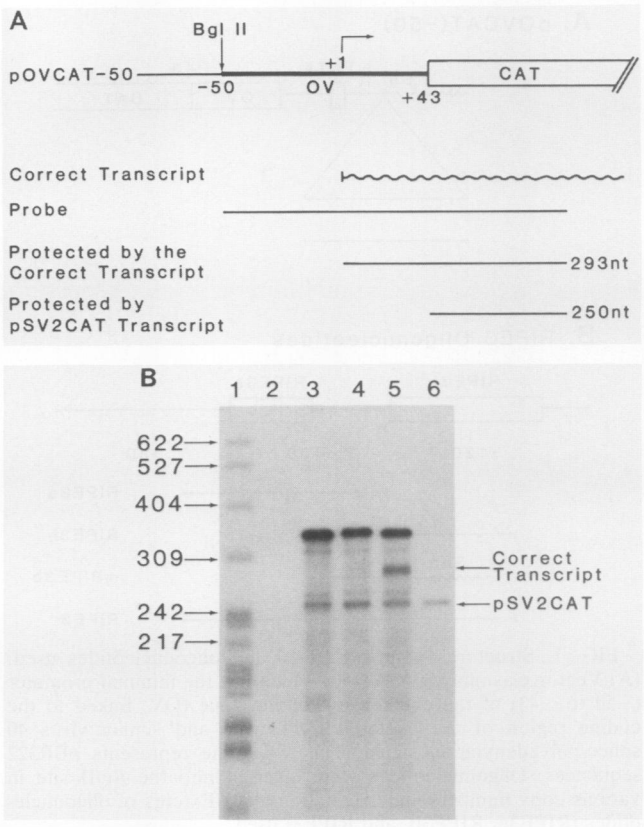

FIG. 4. RNase mapping. (A) Schematic representation of the RNase mapping assay. The wavy line represents the correctly initiated transcript from the ovalbumin (OV) promoter. The hybridization probe spans from the *Bgl*II site to an *Eco*RI site in the CAT gene, as indicated. After RNase treatment, the fragments protected by the correct transcript or the pSV_2CAT transcript are 293 and 250 nucleotides (nt), respectively. (B) Results of the RNase mapping assay. HIT cells were transfected with pSV_2CAT plus pOVCAT-50 (lane 3), pRIPE3(1+) (lane 4), pRIPE3(3+) (lane 5), or pSV_2CAT alone (lane 6). Lane 1, Yeast RNA; lane 2, size markers, with sizes (in nucleotides) shown on the left. The bands protected by the correct transcripts and the pSV_2CAT transcripts are marked on the right.

These constructs were also transfected into non-insulin-producing cells to determine whether RIPE3 activates gene expression in a tissue-specific manner. Figure 5 shows the results of BHK-21 cell transfections. pSV_2CAT was expressed at high levels in BHK cells, indicating that transfection efficiency is high and that the general transcriptional machinery is functional in these cells. As expected, the wild-type insulin plasmid did not display any CAT activity. In contrast to the results of the HIT cell transfections, RIPE3 did not activate gene activity in BHK cells even when it was present in four copies. Similar results were obtained with CV-1, M6, and NIH 3T3 cells (data not shown). Therefore, RIPE3 can confer tissue-specific activation on a heterologous promoter.

To eliminate the possibility that tissue specificity of these constructs came from the minimal promoter of the ovalbumin gene, we inserted an oligonucleotide corresponding to the progesterone response element of the tyrosine aminotransferase gene into pOVCAT-50. The resulting plasmid was transfected into BHK and HIT cells. This construct was active in both BHK and HIT cells (Fig. 5B). Similar results

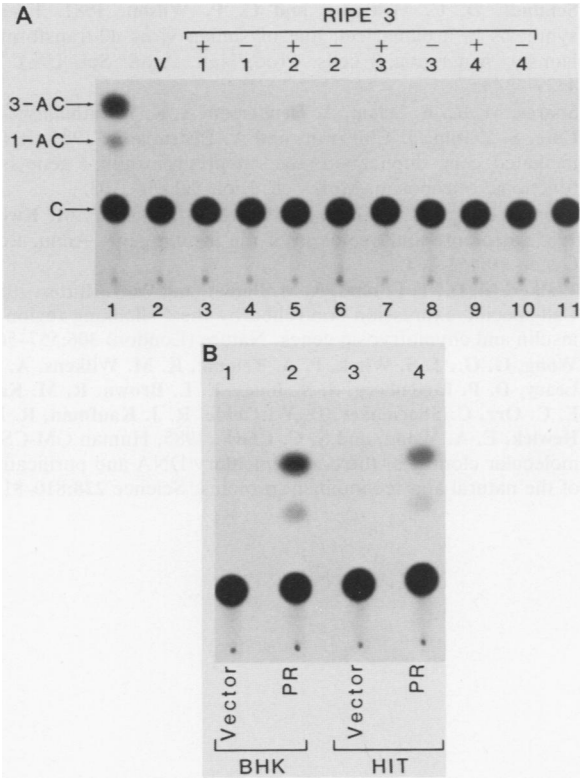

FIG. 5. Activity of RIPE3 in BHK cells. The plasmids used in each experiment in panel A is the same as described in the legend to Fig. 3. In panel B, pOVCAT-50 derivative containing two copies of the progesterone-responsive element was transfected into BHK (lanes 1 and 2) or HIT (lanes 3 and 4) cells. A plasmid (PR) in which progesterone receptor cDNA was cloned into the p91023 expression vector (24) was cotransfected to provide the progesterone receptor in lanes 2 and 4; p91023 (vector) was cotransfected in lanes 1 and 3 as a control.

were observed in CV-1 and M6 cells (data not shown; see reference 2). Therefore, the ovalbumin gene minimal promoter can function in non-insulin-producing cells, and the tissue specificity of the RIPE3–pOVCAT-50 construct is derived from RIPE3.

Two tissue-specific sequence elements have been reported for the other rat insulin gene, rInsI. These sequence elements, the NIR and FAR boxes, are highly homologous to each other (13, 17). The NIR box is located in the same position in the rInsI gene as RIPE3a is in the rInsII gene, and they share a very high sequence similarity, suggesting that their functions may be the same. Indeed, the NIR and FAR boxes can activate the herpes simplex virus thymidine kinase promoter in HIT cells but not in BHK cells (14). Furthermore, the activities of RIPE3a and the NIR box are at a similar level which is much less than that of the full-length insulin enhancer. Therefore, RIPE3a of the rInsII gene may be a functional counterpart of the NIR box of the rInsI gene.

The combined activity of RIPE3a and 3b cannot account for that of RIPE3. Therefore, RIPE3a and RIPE3b may cooperate to achieve the activity of RIPE3, presumably through protein-protein interaction between their cognate *trans*-acting factors. Alternatively, the activity of RIPE3 may come from a sequence element located in the junction of RIPE3a and RIPE3b but not fully covered by either oligonucleotide used in this experiment. However, this latter possibility is less favored for three reasons. First, RIPE3b functions in both orientations and RIPE3a is active only in one orientation, suggesting that two different *trans*-acting factors are involved. Second, the NIR box may be a functional counterpart of RIPE3a, as discussed above. The NIR box per se has been shown to be an independent sequence element that binds specific factors (15). Finally, RIPE3a and 3b oligonucleotides are overlapping by 10 bp. The second possibility predicts that sequences both immediately upstream and downstream from the overlapping region are essential for binding of the putative factor, but it is less likely that a protein will specifically contact bases that are more than 10 nucleotides apart. In addition, preliminary experiments in our laboratory indicate that mutations (4-bp substitution or 5-bp insertion) of the overlapping sequences do not significantly affect RIPE3 activity. Therefore, it is likely that RIPE3a and RIPE3b are separate sequence elements which act cooperatively to confer tissue specificity.

We thank Juan Codina-Salada for oligonucleotide synthesis and Lisa Gamble for typing the manuscript.

This work was supported by Public Health Service grant HD-17379 from the National Institutes of Health to M.-J.T.

LITERATURE CITED

1. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. A. Smith, J. G. Seidman, and K. Struhl (ed.).** 1987. Current protocols in molecular biology. Greene Publishing Associates and Wiley-Intersceince, New York.
2. **Bradshaw, M. S., M.-J. Tsai, and B. W. O'Malley.** 1988. A steroid response element can function in the absence of a distal promoter. Mol. Endocrinol. **2**:1286–1293.
3. **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.
4. **Crowe, D. T., and M.-J. 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.
5. **Edlund, T., M. D. Walker, P. J. Barr, and W. J. Rutter.** 1985. Cell-specific expression of the rat insulin gene: evidence for role of two distinct 5′-flanking elements. Science **230**:912–916.
6. **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.
7. **Fromental, C., M. Kanno, H. Nomiyama, and P. Chambon.** 1988. Cooperativity and hierarchical levels of functional organization in the SV40 enhancer. Cell **54**:943–953.
8. **Giddings, S. J., and L. R. Carnaghi.** 1988. The two nonallelic rat insulin mRNAs and pre-mRNAs are regulated coordinately in vivo. J. Biol. Chem. **263**:3845–3849.
9. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
10. **Hanahan, D.** 1985. Heritable formation of pancreatic β-cell tumors in transgenic mice expressing recombinant insulin/simian virus 40 oncogenes. Nature (London) **315**:15–122.
11. **Hwung, Y.-P., D. T. Crowe, L.-H. Wang, S. Y. Tsai, and M.-J. Tsai.** 1988. The COUP transcription factor binds to an upstream promoter element of the rat insulin II gene. Mol. Cell. Biol. **8**:2070–2077.
12. **Hwung, Y.-P., L.-H. Wang, S. Y. Tsai, and M.-J. Tsai.** 1988. Differential binding of the COUP transcription factor to two different promoters. J. Biol. Chem. **263**:13470–13474.
13. **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 β cells is dependent on two related sequences within the enhancer. Proc. Natl. Acad. Sci. USA **84**:8819–8823.
14. **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.
15. **Ohlsson, H., O. Karlsson, and T. Edlund.** 1988. A β-cell-specific protein binds to the two major regulatory sequences of the insulin gene enhancer. Proc. Natl. Acad. Sci. USA **85**:4228–4231.
16. **Ondek, B., L. Gloss, and W. Herr.** 1988. The SV40 enhancer contains two distinct levels of organization. Nature (London) **333**:40–45.
17. **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.
18. **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.
19. **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.
20. **Santerre, R. F., R. A. Cooke, R. M. D. Crisel, J. D. Sharp, R. J. Schmidt, D. C. Williams, and C. P. Wilson.** 1981. Insulin synthesis in a clonal cell line of simian virus 40-transformed hamster pancreatic β cells. Proc. Natl. Acad. Sci. USA **78**:4339–4343.
21. **Soares, M. B., E. Schon, A. Henderson, S. K. Karathanasis, R. Cate, S. Zeitlin, J. Chirgwin, and A. Efstratiadis.** 1985. RNA-mediated gene duplication: the rat preproinsulin I gene is a functional retroposon. Mol. Cell. Biol. **5**:2090–2103.
22. **Steiner, D. F., S. J. Chan, J. M. Welsh, and S. C. M. Kwok.** 1985. Structure and evolution of the insulin gene. Annu. Rev. Genet. **19**:463–484.
23. **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.
24. **Wong, G. G., J. S. Witek, P. A. Temple, K. M. Wilkens, A. C. Leary, D. P. Luxenberg, S. S. Jones, E. L. Brown, R. M. Kay, E. C. Orr, C. Shoemaker, D. W. Golde, R. J. Kaufman, R. M. Hewick, E. A. Wang, and S. C. Clark.** 1985. Human GM-CSF: molecular cloning of the complementary DNA and purification of the natural and recombinant proteins. Science **228**:810–815.