MOLECULAR AND CELLULAR BIOLOGY, Feb. 1989, p. 823–827
0270-7306/89/020823-05$02.00/0
Copyright © 1989, American Society for Microbiology
Vol. 9, No. 2

Individual Protein-Binding Domains of the Insulin Gene Enhancer Positively Activate β-Cell-Specific Transcription

OLOF KARLSSON,[1] MICHAEL D. WALKER,[2] WILLIAM J. RUTTER,[3] AND THOMAS EDLUND[1]*

Department of Microbiology, University of Umea, S-901 87 Umea, Sweden[1]; Department of Biochemistry, Weizmann Institute of Science, Rehovot 76 100, Israel[2]; and Hormone Research Institute, University of California, San Francisco, California 94143[3]

Received 11 July 1988/Accepted 31 October 1988

A β-cell-specific enhancer is located in the 5′-flanking DNA of the rat insulin 1 gene. Two homologous 8-base-pair sequences in the enhancer (IEB1 and IEB2) significantly stimulated transcription from a heterologous promoter (two- to fourfold) in a cell-specific fashion. When the elements were combined or duplicated, more than 50% of the activity of the intact enhancer was obtained. These two *cis*-acting elements appear to play a dominant role in the positive control of β-cell-specific transcription of the insulin gene.

Transcriptional enhancers are composed of a number of repetitive sequence domains which bind sequence-specific, *trans*-acting proteins (reviewed in references 12 and 25). However, it has proven difficult to analyze the function of the individual domains because of the redundancy of enhancers: individual domains can be deleted or mutated without significantly affecting overall activity. To address this issue, the activity of a single domain as either a single copy or a tandem repeat can be tested. In the case of viral enhancers and inducible enhancers, this leads to regeneration of a significant proportion of the original enhancer activity (3, 13, 24). For cell-specific enhancers, short DNA fragments containing more than one protein-binding site have in several instances been shown to display cell-specific activity when multimerized in the presence of viral enhancer elements (5, 11, 21, 27). However, very few reports have thus far documented the ability of a single binding domain from a cell-specific enhancer to activate cell-specific expression in the absence of other enhancer elements from viral or cellular genes (22). It is therefore not yet clearly established whether cell-specific transcription results from the activity of a single element with intrinsic cell-specific activity or from a combinational effect of several elements.

We have recently shown that the two most mutationally sensitive regions of the insulin enhancer (4, 9, 26) interact with proteins, including an apparently β-cell-specific protein designated IEF1 (19, 20). This prompted us to test whether these two sequences alone possessed cell-specific enhancer activity or whether multiple protein-binding domains are needed for β-cell-specific activity.

Individual IEF1 binding domains of the insulin enhancer can enhance the activity of a heterologous promoter. The two domains designated insulin enhancer boxes 1 and 2 (IEB1 and IEB2), located at positions −104 to −111 and −231 to −238, respectively, have an 8-base-pair (bp) homology, GCCATCTG (Fig. 1). Insulin enhancer sequences (10 bp each) encompassing IEB1 and IEB2 were synthesized as complementary oligonucleotides with a protruding 4-base overhang to permit insertion at *Bam*HI or *Bgl*II restriction enzyme sites. These sites are located upstream of the thymidine kinase (TK) promoter in an expression vector (pOK2, Fig. 1) containing the gene encoding chloramphenicol acetyltransferase (CAT) (6). When these constructions were introduced into HIT cells (an established hamster β-cell line [23]), CAT activity was increased two- to fourfold compared with that observed for the control plasmid pOK2 (Fig. 2, constructions 1, 4, 5, 7, and 8). Positioning of the domains at the *Bam*HI site (immediately adjacent to the TK promoter) or at the *Bgl*II site (22 bp further upstream) gave essentially identical results, indicating that precise positioning was not essential for activity. When two copies of each domain were inserted into the test plasmid, the activities increased to 4.5-fold and 11.4-fold that of the control pOK2 plasmid (Fig. 2, constructions 6 and 9). The combination of these two elements led to 8.5-fold enhancement (Fig. 2, construction 10). A 5.4-fold enhancement could still be observed when these two fragments were simultaneously inserted in the inverted orientation (construction 11).

Enhancing activity of the IEF1 binding site is restricted to insulin-producing cells. To determine whether the IEF1 binding sites act in a cell-specific manner, the activities of the above constructions were analyzed in several cell types which do not produce insulin. No enhancement was observed in hamster fibroblasts (BHK cells) with any of these constructions (Fig. 2), nor was any activity observed in CHO or 3T3 fibroblast lines (data not shown). The IEF1 binding sites were also shown to be active in an independently derived insulin-producing cell line, JHP1 (14) (data not shown). As a control for the sequence specificity of the enhancing activity of these elements, we used a 17-bp synthetic DNA fragment (CCAAT fragment) that is identical to the IEB1 fragment except for the insertion of an A/T pair (Fig. 1). This mutation reduces the activity of the otherwise intact insulin 5′-flanking sequence approximately 10-fold in HIT cells (9). The mutant sequence (GCCAATCTG) interacts in vitro with a protein distinct from IEF1 and present in a variety of cell types (20). The protein may be related to the widely distributed CCAAT-binding transcription factors/nuclear factor I (CTF/NF-I), since the mutant sequence closely resembles a CCAAT upstream promoter element (8). The construction carrying two copies of this mutant sequence did not show any increase in CAT activity compared with the pOK2 control in the two cell lines used (Fig. 2, construct 12). To demonstrate that enhancers can work effectively in the fibroblast lines used, we tested the ability of the Moloney murine sarcoma virus (MSV) enhancer, known to be active in a wide range of cell types (4, 10), to activate the TK promoter of pOK2. As expected, the MSV enhancer showed

* Corresponding author.

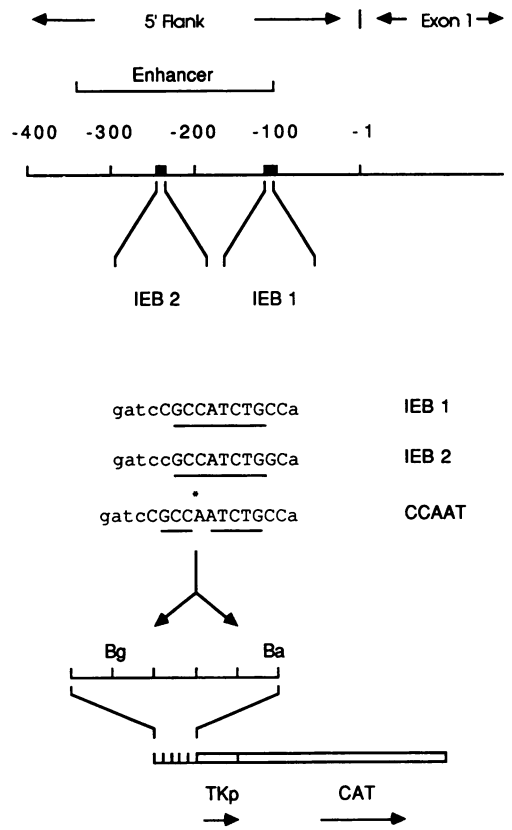
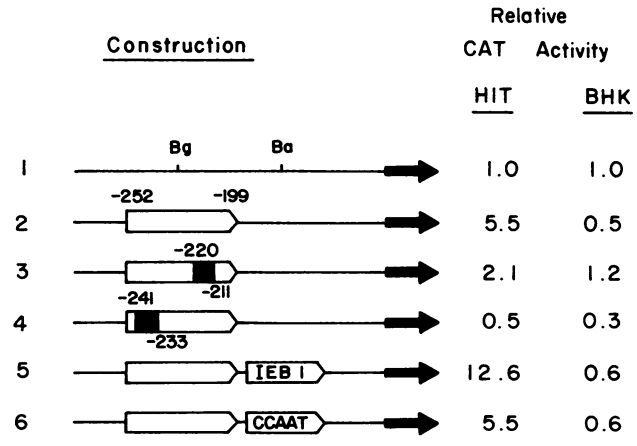


FIG. 1. Construction of plasmids for testing enhancer activity of individual domains. The upper portion of the figure illustrates the position of the enhancer and the IEB1 and IEB2 sequences within the 5'-flanking DNA. The lower portion of the figure shows the sequences of the oligonucleotides used (top strand only). Oligonucleotides encompassing the IEB1 and IEB2 sequences and containing flanking sequences corresponding to the recognition sites for the enzymes *Bam*HI (Ba) and *Bgl*II (Bg) were inserted into pOK2, a CAT test plasmid. pOK2 is a pUC18-based plasmid which contains the *cat* gene linked to simian virus 40 splice and polyadenylation signals, the TK promoter (TKp) region (−109 to +51) (15), and a polylinker sequence containing multiple unique restriction enzyme recognition sites (*Bam*HI, *Bgl*II, *Xba*I, and *Sal*I). The sequences of the synthetic DNA fragments tested for enhancer activity are shown at the top of the figure (top strand only). Capital letters correspond to insulin enhancer sequences, and the underlined nucleotides indicate sequence homology between IEB1 and IEB2. The asterisk indicates the A residue that was inserted into IEB1 to create a CCAAT sequence.



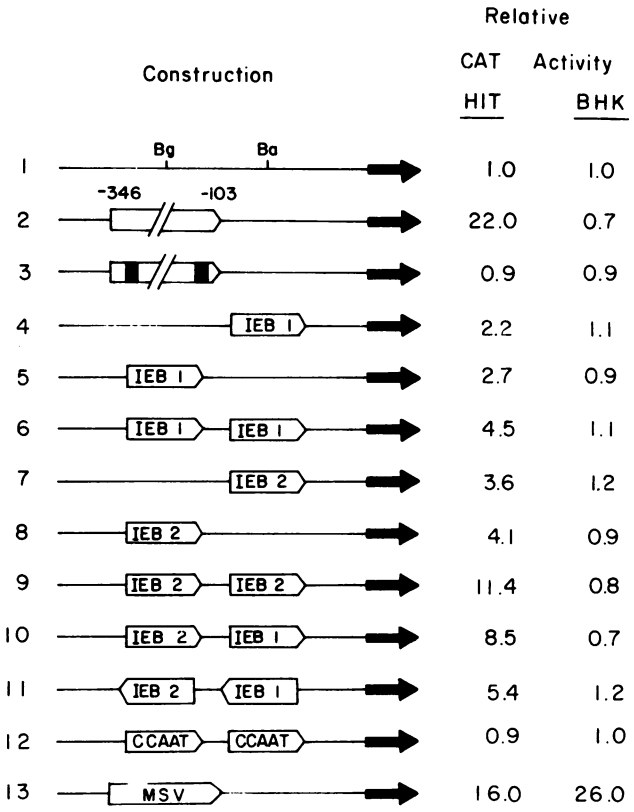


FIG. 2. CAT activities directed by the TK promoter in the presence of insulin and MSV enhancer sequences. The TK promoter is indicated by the solid arrows, and the various enhancer DNA fragments are shown by open arrows. The *Bam*HI (Ba) and *Bgl*II (Bg) sites used for insertion of fragments are indicated. The numbers above the open arrows represent the nucleotide positions of these sequences relative to the start site of transcription. The hatched areas in construction 3 indicate mutated sequences within the insulin enhancer: the sequence from −112 to −104 was changed from 5'-GCCATCTGC-3' to 5'-TAACGAGTA-3', and the sequence from −241 to −233 was changed from 5'-CAGGCCATC-3' to 5'-ACT TAACGA-3'. Activities were determined after transfection of HIT-T15 M2.2.2 (4, 23) and BHK 21 cells and are expressed relative to that of construction 1 (plasmid pOK2). Construction 3 carries the insulin enhancer with mutations in both IEB1 and IEB2. Construction 13 carries the MSV enhancer. Transfection and enzymatic assays were performed with an internally controlled system, based on cotransfection of *cat* plasmids with pRSV-β-gal, a β-galactosidase expression plasmid, as described previously (4, 9). The numbers presented are the means of at least four independent DNA transfections. In all cases, the standard error of the mean was <20% of this value.

FIG. 3. CAT activities directed by the TK promoter in the presence of wild-type and mutant insulin enhancer sequences. Construction 1 corresponds to plasmid pOK2. The numbers in construction 2 correspond to nucleotide positions in the insulin 5'-flanking DNA. The hatched areas and the numbers in constructions 3 and 4 indicate the positions of the mutated sequences in the respective constructions. Sequences from −222 to −211 were mutated from 5'-TTAATAATCTAA-3' to 5'-GGCCGCCGAGCC-3'. Sequences from −241 to −233 were mutated as described in the legend to Fig. 2. Transfection and enzymatic assays were performed as described in the legend to Fig. 2.

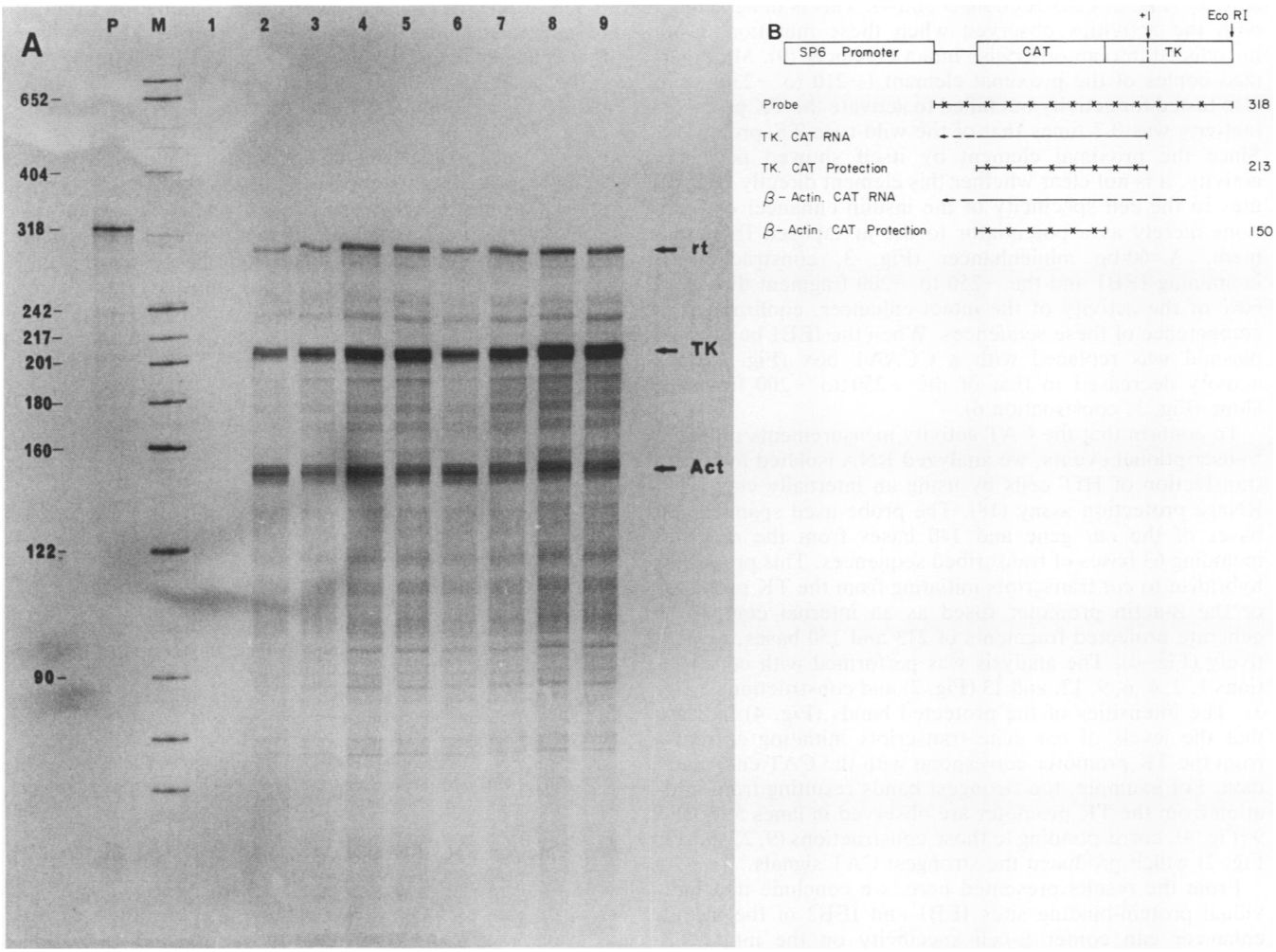


FIG. 4. Analysis of CAT transcripts from transfected HIT cells. (A) Total cellular RNA was extracted from cells following transfection with the following constructions: lane 1, mock transfection; lane 2, construction 1; lane 3, construction 4; lane 4, construction 6; lane 5, construction 9; lane 6, construction 12; lane 7, construction 2; lane 8, construction 2; lane 9, construction 13. In each case, 10 μg of the construction being tested was mixed with 10 μg of the plasmid pβ-CAT (containing 5′-flanking sequences of the rat β-actin [Act] gene [16]) as an internal control. RNA was analyzed by RNase mapping (18) with a probe generated by the action of SP6 polymerase (see below). Lane P contains the probe used for the analysis, and lane M shows radioactively labeled size markers. Sizes are shown on the left. (B) Predicted protected fragments resulting from hybridization of SP6 polymerase-generated probe with *cat* transcripts from transfected cells. The probe is 318 bases long. Transcripts initiating at the TK promoter and β-actin promoter should lead to protected fragments of 213 and 150 nucleotides, respectively. The band at 294 nucleotides probably results from readthrough (rt) from initiation sites in plasmid sequences upstream from the TK promoter.

strong activity in both cell types (26-fold stimulation in BHK cells and 16-fold in HIT cells) (Fig. 2, construction 13). In parallel, a wild-type insulin enhancer fragment and a fragment containing block replacement mutations (9) in both IEB1 and IEB2 were inserted into the *Bgl*II site of plasmid pOK2. The wild-type enhancer stimulated promoter activity 22-fold in HIT cells but not in BHK cells (Fig. 2, construction 2), while the mutated enhancer was inactive in both cell types (Fig. 2, construction 3). Deletion of the TK promoter from the constructions containing a single copy of IEB1 or IEB2 led to complete loss of activity in the two cell lines used (data not shown), demonstrating that these domains of the insulin enhancer are themselves devoid of intrinsic promoter activity and elevate transcription in these experiments by activating the TK promoter.

Additional protein-binding sites potentiate the activity of the IEB2 element. In DNase I protection experiments, most of the −200 to −250 region is protected from cleavage by nuclear extracts from HIT cells (19, 20). In addition to IEB2, this region contains two other mutationally sensitive sequence blocks (−211 to −222 and −223 to −232). Mutation of these sequence blocks results in an approximately threefold reduction of activity (9). The −199 to −252 sequences were assembled by using eight oligonucleotides, and the resulting synthetic DNA fragment was inserted at the *Bgl*II site of plasmid pOK2. This construction (Fig. 3, construction 2) showed enhancement of the TK promoter selectively in HIT cells (5.5-fold). Mutations in sequences −211 to −222 reduced the activity about twofold, whereas mutations in sequences −233 to −241 essentially abolished enhancer

activity (Fig. 2, constructions 3 and 4). This is in agreement with the activities observed when these mutations were introduced into an otherwise intact enhancer (9). Multimerized copies of the proximal element (−210 to −230) were also tested for activity but failed to activate the TK promoter (activity was 0.7 times that of the wild-type TK promoter). Since the proximal element by itself showed negligible activity, it is not clear whether this element directly contributes to the cell specificity of the insulin enhancer or functions merely as a potentiator to the juxtaposed IEB2 element. A 60-bp minienhancer (Fig. 3, construction 5) containing IEB1 and the −250 to −200 fragment displayed 60% of the activity of the intact enhancer, confirming the competence of these sequences. When the IEB1 box of this plasmid was replaced with a CCAAT box (Fig. 1), the activity decreased to that of the −250 to −200 fragment alone (Fig. 3, construction 6).

To confirm that the CAT activity measurements reflected transcriptional events, we analyzed RNA isolated following transfection of HIT cells by using an internally controlled RNase protection assay (18). The probe used spanned 150 bases of the *cat* gene and 140 bases from the *tk* gene, including 63 bases of transcribed sequences. This probe will hybridize to *cat* transcripts initiating from the TK promoter or the β-actin promoter (used as an internal control) to generate protected fragments of 213 and 150 bases, respectively (Fig. 4). The analysis was performed with constructions 1, 2, 4, 6, 9, 12, and 13 (Fig. 2) and construction 2 (Fig. 3). The intensities of the protected bands (Fig. 4) indicate that the levels of *cat* gene transcripts initiating correctly from the TK promoter correspond with the CAT enzymatic data. For example, the strongest bands resulting from initiation from the TK promoter are observed in lanes 5, 8, and 9 (Fig. 4), corresponding to those constructions (9, 2, and 13, Fig. 2) which produced the strongest CAT signals.

From the results presented here, we conclude that individual protein-binding sites IEB1 and IEB2 of the insulin enhancer can confer β-cell specificity on the intact TK promoter. The combination of these two sequences, in the absence of any other sequences from the insulin flank, restores as much as 40% of the activity of the intact insulin enhancer. It has been shown previously that these two *cis*-acting elements are crucial for enhancer activity and that they interact with protein factors (17, 20), including an apparently cell-specific nuclear protein (IEF1 [20]). We therefore infer that these elements are key determinants positively controlling the cell specificity of the insulin gene enhancer. This enhancer contains other elements of lesser but still significant importance for activity which have also been shown to interact with protein factors present in whole-cell (17) and nuclear (20) extracts. One such mutationally sensitive protein-binding region (−201 to −222), which is located close to IEB2, shows no activity by itself but clearly potentiates the activity of IEB2 in insulin-producing cells. Furthermore, the fact that removal of sequences −113 to −199 from the intact enhancer (construction 5, Fig. 3) resulted in less than wild-type activity (about 60%) suggests some involvement of these sequences, consistent with the weak but reproducible reductions observed on analysis of mutations within this region (9). Thus, the rat insulin I gene enhancer contains two major determinants which contribute a significant part of the overall enhancing activity and control the cell type preference and a second set of elements whose precise function is not known. Possibly they stabilize the interaction between IEF1 and other proteins of the transcription apparatus.

These results, together with those obtained in other experimental systems (22), are consistent with the hypothesis that cell specificity results from the activities of intrinsically cell-specific DNA domains and their cognate factors rather than from a combinatorial interaction of several elements, each of which in isolation shows limited specificity. It may be that other protein-binding domains of these two enhancer elements also directly contribute to the cell type preference of these enhancers, but due to their relatively low intrinsic activity it is not possible to detect such activities in the experimental systems used. This may be possible in more physiological test systems, such as transgenic animals, in which the ability of 5′-flanking DNA sequences from the rat insulin gene to direct β-cell-specific expression has been demonstrated (1, 7).

This work was supported by Swedish grants MFR 07149, NFR 4875-107, and STU 84-05425P to T.E., by Public Health Service (PHS) grant GM38817 from the National Institutes of Health (NIH) and a grant from the Israel Academy of Sciences and Humanities to M.D.W., and by PHS grants AM 21344 and GM 28520 from NIH and a grant from the March of Dimes to W.J.R. M.D.W. is supported by a Career Development Award from the Juvenile Diabetes Foundation.

LITERATURE CITED

1. **Alpert, S., D. Hanahan, and G. Teitelman.** 1988. Hybrid insulin genes reveal a developmental lineage for pancreatic endocrine cells and imply a relationship with neurons. Cell **53**:295–308.
2. **Chen, C., and H. Okayama.** 1987. High-efficiency transformation of mammalian cells by plasmid DNA. Mol. Cell. Biol. **7**:2745–2752.
3. **Chiu, R., M. Imagawa, R. J. Imbra, J. R. Backoven, and M. Karin.** 1987. Multiple *cis*- and *trans*-acting elements mediate the transcriptional response to phorbol esters. Nature (London) **329**:648–651.
4. **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.
5. **Gerster, T., P. Matthias, M. Thali, J. Jiricny, and W. Schaffner.** 1987. Cell type specificity elements of the immunoglobulin heavy chain gene enhancer. EMBO J. **6**:1323–1330.
6. **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.
7. **Hanahan, D.** 1985. Heritable formation of pancreatic β-cell tumors in transgenic mice expressing recombinant insulin/simian virus 40 oncogene. Nature (London) **315**:115–122.
8. **Jones, K. A., J. T. Kadonaga, P. J. Rosenfeld, T. J. Kelly, and R. Tjian.** 1987. A cellular DNA-binding protein that activates enkaryotic transcription and DNA replication. Cell **48**:79–89.
9. **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.
10. **Laimins, L. A., G. Khoury, C. Gorman, B. Howard, and P. Gruss.** 1982. Host-specific activation of transcription by tandem repeats from simian virus 40 and Moloney murine sarcoma virus. Proc. Natl. Acad. Sci. USA **79**:6453–6457.
11. **Lenardo, M., J. W. Pierce, and D. Baltimore.** 1987. Protein-binding sites in immunoglobulin gene enhancers determine transcriptional activity and inducibility. Science **236**:1573–1577.
12. **Maniatis, T., S. Goodbourn, and J. A. Fischer.** 1987. Regulation of inducible and tissue-specific gene expression. Science **236**: 1237–1245.
13. **Martinez, E., F. Givel, and W. Wahli.** 1987. The estrogen-response element as an inducible enhancer: DNA sequence requirements and conversion to a glucocorticoid responsive element. EMBO J. **6**:3719–3727.
14. **Matsuba, I., B. Marner, J. Nerup, and A. Lernmark.** 1987. Detection of islet cell autoantibodies in newly diagnosed dia-

betic patients using insulin-producing Syrian hamster cells. Diabetes Res. **4**:103–107.
15. **McKnight, S. L., and R. Kingsbury.** 1982. Transcriptional control signals of a eukaryotic protein coding gene. Science **217**:316–324.
16. **Melloul, D., B. Aloni, J. Calvo, D. Yaffe, and U. Nudel.** 1984. Developmentally regulated expression of chimeric genes containing muscle actin DNA sequences in transfected myogenic cells. EMBO J. **3**:983–990.
17. **Moss, L., 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. **Nir, U., M. D. Walker, and W. J. Rutter.** 1986. Regulation of rat insulin I gene expression: evidence for negative regulation in non-pancreatic cells. Proc. Natl. Acad. Sci. USA **83**:3180–3184.
19. **Ohlsson, H., and T. Edlund.** 1986. Sequence-specific interactions of nuclear factors with the insulin gene enhancer. Cell **45**:35–44.
20. **Ohlsson, H., O. Karlsson, and T. Edlund.** 1988. A beta-cell-specific protein binding to two related DNA sequences in the insulin gene enhancer. Proc. Natl. Acad. Sci. USA **85**:4228–4231.
21. **Parslow, T., S. D. Jones, B. Bond, and K. R. Yamamoto.** 1987. The immunoglobulin octanucleotide: independent activity and selective interaction with enhancers. Science **235**:1498–1501.
22. **Pierce, J. W., M. Lenardo, and D. Baltimore.** 1988. Oligonucleotide that binds nuclear factor NF-κB acts as a lymphoid-specific and inducible enhancer element. Proc. Natl. Acad. Sci. USA **85**:1482–1486.
23. **Santerre, R. F., R. A. Cook, 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 cells. Proc. Natl. Acad. Sci. USA **78**: 4339–4343.
24. **Schirm, S., J. Jiricny, and W. Schaffner.** 1987. The SV40 enhancer can be dissected into multiple elements each with a different cell type specificity. Genes Dev. **1**:65–74.
25. **Serfling, E., M. Jasin, and W. Schaffner.** 1985. Enhancers and eukaryotic gene transcription. Trends Genet. **1**:224–230.
26. **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.
27. **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.
28. **Yanisch-Perron, C., J. Veira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.