

The LIM Domain Homeobox Gene *isl-1* Is a Positive Regulator of Islet Cell-specific Proglucagon Gene Transcription*

(Received for publication, February 3, 1995, and in revised form March 27, 1995)

Min Wang and Daniel J. Drucker‡

From the The Banting and Best Diabetes Centre and the Departments of Clinical Biochemistry, Medical Genetics, and Medicine, The Toronto Hospital, University of Toronto, Toronto, Canada M5G 2C4

The LIM domain homeobox gene islet 1 (*isl-1*) is expressed in the embryonic nervous system and may be an early marker of motor neuron specification. *isl-1* is expressed in all 4 islet cell types, but a role for *isl-1* in the regulation of insulin gene expression has not been demonstrated, and the genetic targets for *isl-1* in the pancreas remain unknown. We show here that the proximal rat proglucagon gene promoter binds an amino-terminally truncated Trp-E-*isl-1* fusion protein that lacks the LIM domains. The proglucagon gene promoter also binds full-length *in vitro* translated *isl-1* containing the intact LIM domains. *isl-1* antisera detects binding of proglucagon gene sequences to *isl-1* present in a slowly-migrating complex in nuclear extracts from InR1-G9 islet cells. The transcriptional properties of the proglucagon gene promoter sequences that bind *isl-1* (designated Ga, Gb, and Gc) were assessed after transfection of reporter genes into wild type and *isl-1*-antisense (*isl-1*(AS)) InR1-G9 islet cells. The proximal proglucagon gene (Ga) promoter sequence reduced TK-CAT activity by ~50%, but no change in the activity of the Ga-TK-CAT plasmid was seen after transfection of *isl-1*(AS) InR1-G9 cells. In contrast, the Gb/Gc sites activated transcription 2–3-fold in wild type InR1-G9 cells, and the *isl-1*-dependent activation of gene transcription through the Gb/Gc element was eliminated following transfection of *isl-1*(AS) InR1-G9 cells. These data demonstrate that the LIM domain homeobox gene *isl-1* is not constrained from DNA binding by its LIM domains and 2) functions as a positive regulator of proglucagon gene transcription in the endocrine pancreas.

Glucagon and the glucagon-like peptides are encoded within a common precursor, proglucagon, that is expressed in a highly tissue-specific manner in the brain, intestine, and endocrine pancreas (1). Glucagon, secreted from the pancreatic A cell, is a key regulator of carbohydrate, lipid, and protein metabolism, and excess glucagon secretion contributes to the metabolic derangements characteristic of diabetes mellitus. Recent studies have implicated glucagon-like peptide-1 as an important mediator of glucose-dependent insulin secretion from the pancreatic islet (2). Accordingly, understanding the factors important for the regulation of proglucagon gene expression is of

* This work was supported in part by Grant JDF 192425 from Juvenile Diabetes Foundation International and Grant MRC MT10903 from the Medical Research Council of Canada (MRC). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Supported by a Scientist award from the MRC. To whom correspondence should be addressed: Toronto General Hospital, 200 Elizabeth St. CCRW3–838, Toronto, Canada M5G 2C4. Tel.: 416-340-4125; Fax: 416-978-4108.

considerable interest and importance.

The isolation of the genes and cDNAs encoding proglucagon has demonstrated that proglucagon gene transcription originates from a single transcription start site and gives rise to an identical proglucagon mRNA transcript in each tissue (3). Nevertheless, gene transfer and transgenic experiments have demonstrated that the control of proglucagon gene transcription is highly tissue-specific, with different regions of the proglucagon gene promoter identified as necessary for transcriptional activation in the pancreas and intestine (4–7). A series of deletional and mutational analyses have identified several cis-acting DNA sequences in the proximal proglucagon gene promoter that appear to be important for islet cell-specific gene transcription (8). Two of these regions, originally designated G2 and G3, have been shown to act as islet cell-specific enhancers, whereas a more proximal region, G1, displays islet cell-specific promoter activity (8, 9). These DNA sequences have been shown to bind distinct proteins, including several that may be unique to islet cell lines (10, 11). Although these proteins are candidate transcriptional regulators that likely contribute to islet cell-specific proglucagon gene transcription, the identity of these factors has not yet been elucidated.

Recent studies suggest a role for homeobox genes in the control of islet cell-specific gene transcription. Two novel homeobox proteins isolated from islet cell cDNA libraries by expression cloning, Lmx-1 and Cdx-3, bind to sequences in the rat insulin gene promoter and activate insulin gene transcription (12). Degenerate oligonucleotides complementary to conserved sequences of the homeobox were used in polymerase chain reaction experiments to isolate over a dozen distinct transcripts encoding homeobox sequences from islet RNA (13, 14); however, the functional importance of these homeobox proteins in the regulation of islet gene expression has not been established.

A cDNA encoding the homeobox gene *isl-1* was originally isolated by screening a rat islet cell (RIN cell) library with a probe from the insulin gene enhancer (15). The homeobox gene *isl-1* is expressed in all four principal cell types of the endocrine pancreas (16, 17), however a role for *isl-1* in the regulation of insulin gene transcription has not been established (12). Subsequent experiments localized *isl-1* to neurons in the central and peripheral nervous system (16, 17), and developmental analysis of *isl-1* in the embryonic chick spinal cord has shown that *isl-1* is regulated by inductive signals and may serve as a useful early marker of motor neuron differentiation (18). Additional evidence for the biological importance of *isl-1* derives from analyses of *isl-1* genes in different species that demonstrate 100% conservation of the *isl-1* amino acid sequence in the human, hamster, and rat *isl-1* genes (19).

The proglucagon gene proximal promoter contains several AT-rich sequences that are candidate binding sites for homeobox transcription factors (8). Furthermore, the highly re-

TABLE I
TAAT sequences in the proximal promoter regions of the insulin, glucagon, and amylin genes

The location (relative to the transcription start site) and surrounding nucleotide sequences of the TAAT motifs (underlined) from the rat insulin (*rIns I*), human amylin (*hAmylin*), and rat proglucagon (*rGlu*) gene promoters are shown. Mutations in the nucleotide sequences are designated M, and the specific mutated nucleotides are shown in boldface.

Gene	Name	Location	Sequence
<i>rIns I</i>	E2	-230 to -208	GCCCCCTTGTAAATA <u>ATCTAATT</u> A
	E2M		GCCCCCTTGT <u>GAACAAGCTGACTA</u>
<i>rIns I</i>	P1	-85 to -53	GCCC <u>TTAATGGGCCAACGGCAA</u>
<i>hAmylin</i>	AMY	-156 to -137	GAG <u>TAAATGTAATAATGACC</u>
<i>rGlu</i>	Ga	-60 to -36	GCG <u>TAATATCTGCAAGGCTAAACAG</u>
	GaM1		GCG <u>TAATATCTGCAAGGCGCAACAGA</u>
	GaM2		GCG <u>GCATATCTGCAAGGCTAAACAGA</u>
	GaM3		GCG <u>TAATGCCTGCAAGGCTAAACAGA</u>
	Gc	-73 to -49	ATTTATATTGTCAGCG <u>TAATATCTG</u>
	Gb	-95 to -69	CCCC <u>CATTATTACAGATGAGAAATT</u> TA

stricted cell-specific expression of the proglucagon gene suggests that a combination of transcriptional regulators expressed in a tissue-specific manner likely contributes to islet cell-specific expression of the proglucagon gene in the islet A cell. To ascertain whether *isl-1* may be a candidate transcription factor that controls proglucagon gene expression in the islets, we have examined the importance of *isl-1* in the regulation of proglucagon gene transcription.

MATERIALS AND METHODS

EMSA¹ Experiments Using a TrpE-*isl-1* Fusion Protein—A 702-base pair polymerase chain reaction fragment of *isl-1* cDNA (hamster, rat, and human *isl-1* amino acid sequences are 100% identical (19)) containing the complete homeodomain and 3'-sequences from amino acid 118 to the stop codon was sequenced and then inserted in the TrpE expression vector pET 11 to generate the plasmid TrpE-*isl-1*. The standard binding reaction contained the following components in a final volume of 18 μ l: 10 mM HEPES (pH 7.8), 75 mM KCl, 2.5 mM MgCl₂, 0.1 mM EDTA, 1 mM dithiothreitol, 3% Ficoll, 1 mg/ml bovine serum albumin, 500 ng (when using nuclear extract, see Figs. 2–4) or 1 μ g (for TrpE-*isl-1* fusion protein) of poly(dI-dC), 10,000 cpm of end-labeled, double-stranded oligonucleotide, and 12 μ g of crude nuclear extract (Figs. 2–4) or 2 μ g of TrpE-*isl-1* fusion protein extract. The mixture was incubated for 20 min at 30 °C and loaded directly onto a 5% nondenaturing polyacrylamide gel. For experiments with antisera (Figs. 2–4), 5 μ l *isl-1*-specific (17) or preimmune antisera was mixed with the reticulocyte lysate or nuclear extract immediately before adding the premixture. Nuclear extracts from InR1-G9, BHK, and COS7 cells were prepared as described (20). The EMSA was performed essentially as in Ref. 21. The EMSA experiments employed double-stranded synthetic oligonucleotides as probes, corresponding to the sequences shown in Table I, with an additional 5 bases comprising a 5'-*Bam*HI overhang (top strand) and a 5 base 5'-*Bgl*II overhang (bottom strand) for end-labeling. The nonspecific competitor DNA was a double-stranded oligonucleotide containing a GATA binding site from the human cardiac a-myosin heavy chain gene (5'-TTCAACGTGCAGCCGGAGATAAGGCCAG-GCCGAA-3').

In Vitro Translation—The full-length rat *isl-1* cDNA was transcribed from linearized plasmid templates with T7 RNA polymerase, Riboprobe II Core kit, and 5 mM m⁷G(5')ppp(5')G for capping synthesized transcripts in a final volume of 50 μ l. RNase-free DNase I was added to the solution to remove the DNA template for 15 min at 37 °C. The quality of transcript synthesis was checked on a 1% agarose gel. Approximately 5 μ g of RNA were added to a 50- μ l *in vitro* translation mixture including nuclease-treated rabbit reticulocyte lysate in accordance with the manufacturer's instructions. A control translation mix with no added RNA was treated identically. Identical translation reactions using [³⁵S]-labeled methionine were performed in parallel with the *isl-1* template to verify that equivalent amounts of protein were synthesized in each experiment. The labeled translation reactions were analyzed on a 10% acrylamide gel.

Analysis and Generation of Antisense Cell Lines—Isolation of RNA and Northern blot analysis was carried out as described. Four different

isl-1 antisense expression plasmids *isl-1*(AS)A–D contained various lengths of *isl-1* 5'-untranslated, intron and translated sequences. Plasmid AS-*isl-1*A contained 440-bp rat *isl-1* genomic sequences 5' to the translation start site and 30-bp 5'-coding sequences. This 470-bp *isl-1* fragment was subcloned into the *Eco*RI and *Hind*III sites of plasmid pSR1neo in the antisense orientation. Plasmid AS-*isl-1*B contained 100-bp rat *isl-1* genomic sequences 5' to the translation start site and 30-bp 5'-coding sequences (total 130 bp) subcloned into the same restriction sites in pSR1neo. Plasmid AS-*isl-1*D contained hamster full-length *isl-1* coding sequences (19). This 1060-bp fragment was subcloned into the *Hind*III and *Bam*HI sites of the mammalian expression vector pSR1neo in the antisense orientation. Plasmid AS-*isl-1*C was produced by ligating a 360-bp *Eco*RI fragment of AS-*isl-1*D (containing sequences from the translation start site to the *Eco*RI site of the *isl-1* cDNA) into pSR1neo (22). All of the antisense constructs were confirmed by DNA sequencing. G418-resistant InR1-G9 clones transfected with the *isl-1* antisense expression plasmids were individually expanded, and the levels of *isl-1* RNA and protein were assessed by Northern and Western blotting as described previously (19).

Construction of Reporter Plasmids and Cell Transfections—InR1-G9 cells were transfected in suspension by the DEAE-dextran method (6). Minihancer multimers were constructed from the appropriate oligonucleotides (Table I) synthesized with restriction enzyme *Bam*HI and *Bgl*II recognition sites on opposite ends. The ligated oligonucleotides were inserted into the *Bam*HI site of the PUTKAT vector adjacent to the thymidine kinase promoter (23). All plasmids were sequenced to verify copy number and orientation. Cells were harvested 48 h after transfection and analyzed for chloramphenicol acetyltransferase activity as described (6). Transfection efficiencies were monitored by cotransfected the plasmid Rous Sarcoma virus β -galactosidase and measuring the β -galactosidase activity in each sample. CAT activity was normalized for variations in protein concentration and transfection efficiency.

RESULTS

Analysis of the sequence of the rat proglucagon gene promoter (8) identified 2 AT-rich sequences containing TAAT motifs in the first 100 bp upstream of the transcription start site. To determine if *isl-1* binds to the proglucagon gene TAAT motifs, we prepared a TrpE-*isl* fusion protein lacking the *isl-1* LIM domains, since previous studies have shown that the LIM domains inhibit homeodomain binding of DNA sequences *in vitro* (24, 25). The amino-terminally truncated *isl-1* fusion protein was incubated with synthetic oligonucleotides containing TAAT motifs from the proximal (first 100 bp) rat proglucagon gene promoter (Table I). All 3 proglucagon gene sequences (Ga/Gb/Gc) bound the TrpE-*isl-1* fusion protein in an electrophoretic mobility shift assay (Fig. 1). Competition experiments were carried out using the Ga/Gb/Gc sequences as well as oligonucleotides containing TAAT sequences from the insulin and amylin promoters (E2, P1, and AMY) (Table I). A 200-fold molar excess of unlabeled Ga markedly diminished the binding of *isl-1* to the Ga probe (Fig. 1). In contrast, 1000-fold molar excess Gb was much less effective in competing for Ga binding, suggesting that the core TAAT motif alone (present in Ga/Gb/Gc) is not the only determinant for *isl-1* binding. Furthermore,

¹ The abbreviations used are: EMSA, electrophoretic mobility shift assay; CAT, chloramphenicol acetyltransferase; TK, thymidine kinase; AS, antisense; bp, base pair(s).

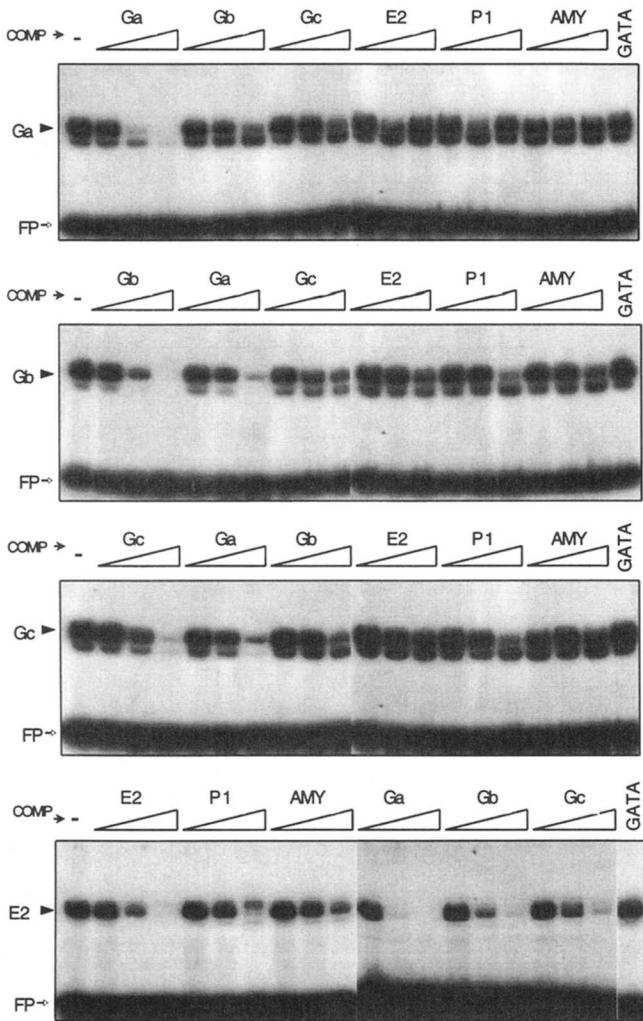


FIG. 1. EMSA analysis of *isl-1* binding to the proglucagon gene Ga, Gb, Gc, and insulin gene E2 sequences. Increasing amounts (20, 200, or 1000-fold molar excess) of designated unlabeled oligonucleotides (shown at the top of each figure; see Table I) were used to compete for *isl-1* binding. COMP, competitor DNA sequence; FP, free probe. 10,000 cpm of end-labeled, double-stranded oligonucleotide and 2 µg of unpurified bacterial extract containing the TrpE-*isl-1* fusion protein were used in each reaction.

the Gc sequence, which contains the same TAAT motif as Ga but different 5'- and 3'-flanking nucleotide sequences was much less effective in displacing Ga binding, providing further evidence that the nucleotide sequences surrounding the TAAT motif determine binding affinity (Fig. 1). In contrast, no displacement of Ga binding was detected with a 1000-fold molar excess of the TAAT-containing amylin or insulin E2 and P1 elements (Fig. 1). Although unlabeled excess Gb effectively competed for *isl-1* binding to Gb, both Ga and Gc also diminished specific *isl-1* binding to Gb, although not as effectively as Gb alone (Fig. 1). In contrast, a 1000-fold molar excess of unlabeled E2, P1, and AMY probes was necessary to achieve a reduction of specific *isl-1* binding to the Gb probe. The proglucagon gene Gc sequence was also effective in binding *isl-1*, and *isl-1* binding to Gc was displaced by unlabeled Gc as well as by a 200-fold molar excess of Ga. Competition for *isl-1* binding to Gc was also observed with a 1000-fold molar excess of Gb and P1 sequences. In contrast, no displacement of *isl-1* binding to Gc was detected with a 1000-fold molar excess of the E2, AMY, or GATA oligonucleotides (Fig. 1). The *isl-1* homeodomain also bound the E2 insulin gene probe, but the binding to E2 was more effectively diminished in competition experiments using

the proglucagon gene Ga/Gb/Gc sequences as competitors, compared with the competition observed with unlabeled E2 alone (Fig. 1).

The observation that the *isl-1* homeodomain (without the LIM domains) binds sequences in the proximal proglucagon gene promoter suggested that *isl-1* is a candidate regulator of proglucagon gene transcription. To determine if an *isl-1* protein including the LIM domains could also bind to the proglucagon gene promoter, we prepared full-length *in vitro* translated *isl-1* for use in EMSA experiments. Contrary to previous reports (that the *isl-1* LIM domain sequences inhibit *isl-1* binding (24, 26)), we detected a specific DNA-protein complex using *in vitro* translated *isl-1* and the Ga probe (Fig. 2A). Addition of *isl-1* antisera to the EMSA reaction eliminated the formation of the *isl-1*-Ga complex. To detect evidence for *isl-1* binding to the Ga sequence in islet cells, we incubated nuclear extracts from the InR1-G9 glucagon-producing islet cell line with the proglucagon gene Ga sequence, producing several distinct DNA-protein complexes (Fig. 2A); addition of anti-*isl-1* antisera to the EMSA reaction resulted in a clear supershift of a high molecular weight complex (Ab + *isl-1*; Fig. 2A, lane 4). Addition of excess *in vitro* translated *isl-1* to the InR1-G9 extract eliminated the formation of both the high molecular weight complex and the complex designated B, and *isl-1* antisera attenuated the formation of the *in vitro* translated *isl-1*/Ga complex in the presence of InR1-G9 extract, resulting in a slight supershift of this lower band (*isl-1* + Ab; Fig. 2A).

A supershift of the high molecular weight *isl-1* complex in InR1-G9 extract was also seen with the Gb and Gc probes (data not shown). The high molecular weight *isl-1* complex was markedly diminished after incubation with a 500-fold molar excess of unlabeled competitor sequences from islet hormone gene promoters, including the Ga, Gb, Gc, E2, AMY, and P1 sequences (Fig. 2B). In contrast, no significant diminution of complex formation was detected after incubation with excess GATA oligonucleotide. Addition of *isl-1* antisera also resulted in increased intensity of complex A. The formation of this complex was attenuated in the presence of excess competitor Ga sequences but was not comparably diminished after competition with the other TAAT-containing oligonucleotides (Fig. 2B).

To determine the DNA sequence requirements for *isl-1* binding to the proglucagon Ga sequence, we synthesized three Ga mutants (Table I) and tested their *isl-1*-binding properties in the EMSA experiment using InR1-G9 nuclear extracts. Mutant GaM1, which contains an intact TAAT motif but two nucleotide substitutions in the nucleotides 12 and 13 bp 3' to the TAAT site also formed a high molecular weight complex that was supershifted with *isl-1* antisera (Fig. 2C). In contrast, the GaM2 sequence with a disrupted TAAT motif did not form the high molecular weight *isl-1* containing complex with InR1-G9 extract (Fig. 2C). Furthermore, mutation GaM3 (that contains an intact TAAT sequence but an AT-GC mutation of the two nucleotides immediately 3' to the TAAT site) also failed to form the high molecular weight *isl-1* complex. These observations suggest that although the TAAT site appears to be important for *isl-1* binding to Ga, nucleotides that flank the TAAT motif are also critical for the formation of the proglucagon gene-*isl-1* complex. Moreover, although the specific mutations in GaM2 and GaM3 eliminated *isl-1* binding (Fig. 2C), no effect on the formation of complex A (which is specifically competed by excess Ga; Fig. 2B) was seen in the same experiment (Fig. 2C). These data suggest that the nucleotide determinants of *isl-1* binding to Ga are highly specific, since the same Ga mutations did not affect protein binding and complex A formation.

To ascertain whether the proglucagon gene sequences that

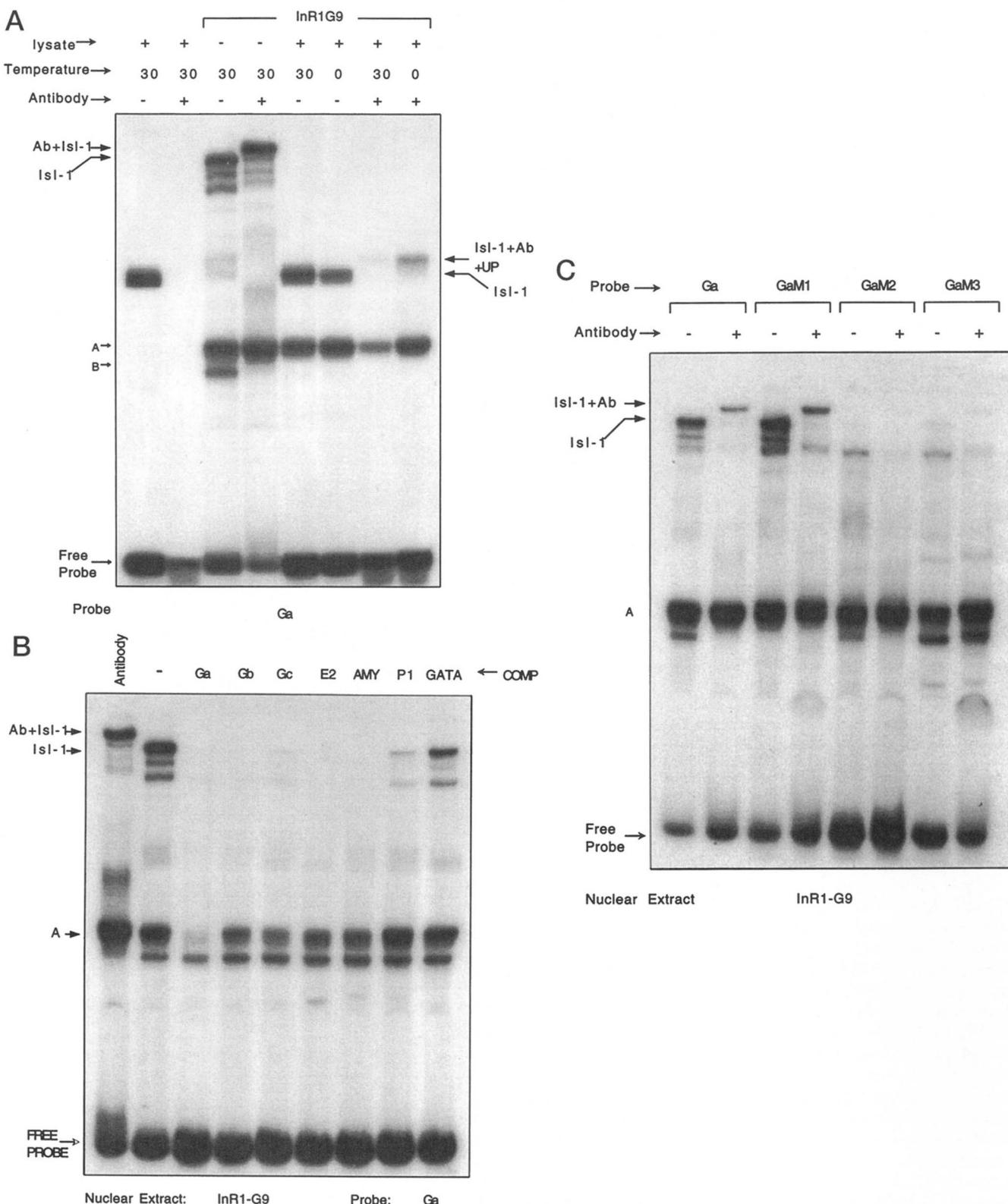


FIG. 2. The proglucagon gene sequences bind full-length *in vitro* translated *isl-1* as well as an *isl-1*-immunoreactive protein present in InR1-G9 islet cell nuclear extracts. Panel A, lysate products (from 5–10% of each reaction) were used for EMSA reactions in the absence or presence of nuclear extracts, and *isl-1* antisera (antibody). The incubations were carried out at either 30 °C or 0 °C as shown. A and B designate as yet unidentified DNA-protein complexes. Panel B, competition for *isl-1*/proglucagon gene-Ga complex formation in InR1-G9 extract in the presence of specific and nonspecific (500-fold molar excess) competitor sequences. The competitor oligonucleotides are shown in Table I. Ab + *isl-1* designates the supershifted complex seen with immune but not preimmune *isl-1* antisera. A refers to a DNA-protein complex not yet definitively identified. Panel C, effect of mutations in the proglucagon gene Ga sequence on *isl-1* binding and complex formation in InR1-G9 nuclear extract. The oligonucleotides containing specific mutations, shown in Table I, were used in EMSA analyses with InR1-G9 nuclear extract, in the presence (+) and absence (-) of *isl-1*-specific antisera. A denotes a specific DNA-protein complex, as seen in panels A and B.

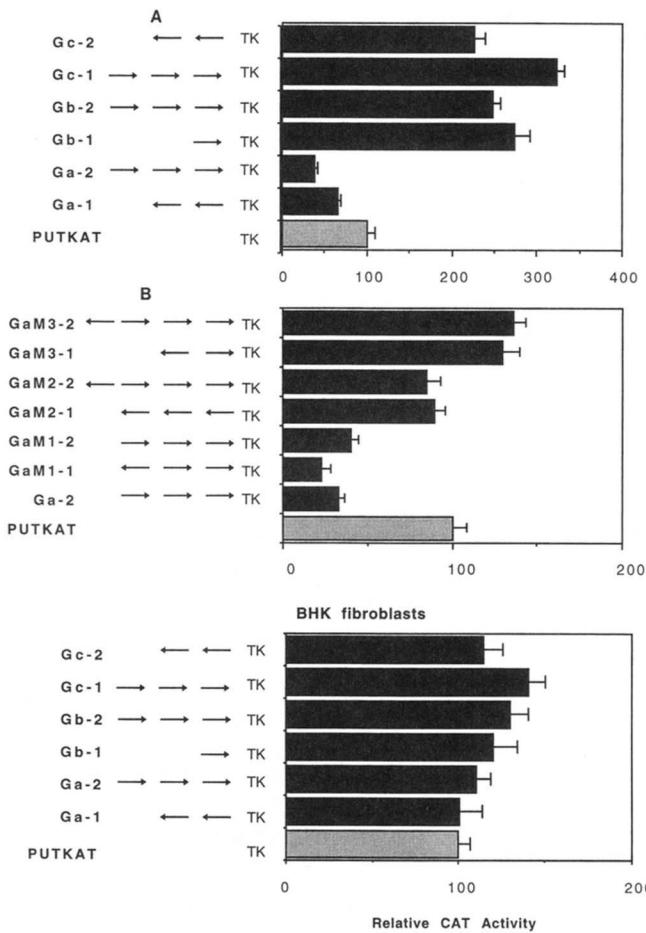


FIG. 3. Transcriptional properties of the proglucagon gene enhancer sequences in InR1-G9 islet cells and BHK fibroblasts. Minienhancer multimers were constructed from the appropriate wild type (*A*) and mutant (*B*) oligonucleotides (Table I), and ligated into the *Bam*H site of the PUTKAT vector (23), immediately 5'-to-the TK promoter. All constructs were sequenced to verify the correct copy number and orientation (arrows) of the inserted oligonucleotides. InR1-G9 (*A* and *B*) and BHK cell transfections (each plasmid in triplicate) and CAT assays were carried out, as described previously (6). The data shown represent the mean \pm S.E. of three separate experiments.

bind *isl-1* are functionally important for proglucagon gene transcription, we ligated the Ga/Gb/Gc sequences adjacent to the thymidine kinase promoter, and the transcriptional activity of the different TK-CAT plasmids was assessed following transfection of InR1-G9 cells. The Ga sequence repressed the basal activity of the TK promoter to 45% of control values (Fig. 3). The GaM1 mutation, which displayed intact *isl-1* binding in EMSA experiments (Fig. 2C), also repressed transcriptional activity; however, no repression was seen with either the GaM2 or GaM3 mutants that failed to bind *isl-1* (Fig. 3). In contrast (to the repression of transcriptional activity observed with the wild type Ga sequence) the Gb and Gc proglucagon gene sequences strongly activated transcription (2–3-fold) from the TK promoter in InR1-G9 cells (Fig. 3A) but not in BHK fibroblasts (Fig. 3), NIH3T3 fibroblasts, or JEG choriocarcinoma cells (data not shown).

Since the EMSA experiments identified more than one DNA-protein complex that formed with the Ga/Gb/Gc sequences, we wished to determine the specific contribution of *isl-1* to the control of proglucagon transcription as mediated by the Ga/Gb/Gc elements. *isl-1* antisense expression vectors were constructed and transfected into wild type InR1-G9 cells. After selection with G418, surviving InR1-G9 clones were individually expanded, and *isl-1* mRNA transcripts in each clone were

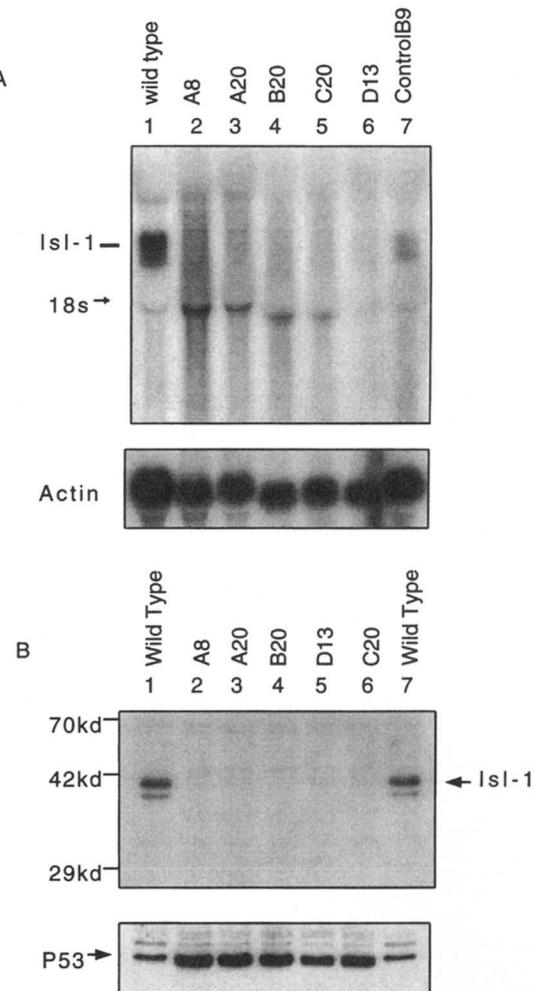


FIG. 4. A, inhibition of *isl-1* gene expression in InR1-G9 *isl-1*(AS) cell lines. The antisense clones shown in lanes 2–6 were derived from transfection with different *isl-1* antisense constructs *A* (lanes 2 and 3), *B* (lane 4), *C* (lane 5), and *D* (lane 6) that contained various amounts of *isl-1* sequences in the 3'-5' orientation. Control B9 represents a G418-resistant InR1-G9 clone transfected with the psR1neo expression vector alone. The Northern blot shown in panel *A* was hybridized with an *isl-1* DNA probe containing mostly 5'-untranslated sequences. The relative migration position of the 18 S ribosomal RNA is indicated with an arrow. *B*, Western blot analysis of immunoreactive *isl-1* in wild type and InR1-G9(AS) cells. 40 μ g of nuclear protein from wild type and antisense InR1-G9 cells was electrophoresed on a 10% SDS-PAGE gel and transferred to a nitrocellulose membrane for hybridization with *isl-1*-specific antisera (17). A 39-kDa *isl-1*-immunoreactive band was detected in the wild type InR1-G9 cells (and G418-resistant controls, not shown) but was greatly reduced in InR1-G9(AS) cells. The relative migration positions of molecular size markers are indicated. The blot was rehybridized with a monoclonal antisera (AB-3) against the nuclear protein p53 (Oncogene Science, Uniondale, NY).

analyzed by Northern blotting. Several InR1-G9 *isl-1*(AS) clones were obtained that contained markedly reduced levels of *isl-1* mRNA transcripts (Fig. 4A). To ascertain whether these clones also contained reduced levels of immunoreactive *isl-1*, nuclear extracts were prepared and analyzed by Western blotting (Fig. 4B). These experiments demonstrated that the InR1-G9 *isl-1*(AS) clones also contained markedly reduced levels of immunoreactive *isl-1*. We next transfected the TK-CAT plasmids containing the Ga/Gb/Gc sequences into the InR1-G9 *isl-1*(AS) cells (Fig. 5). No change in the relative transcriptional activity (compared with the results observed in wild type cells) of the Ga-TK-CAT plasmids was detected following transfection of the *isl-1*-depleted InR1-G9 *isl-1*(AS) cells. In contrast, the *isl-1*-dependent transcriptional activation conferred by the

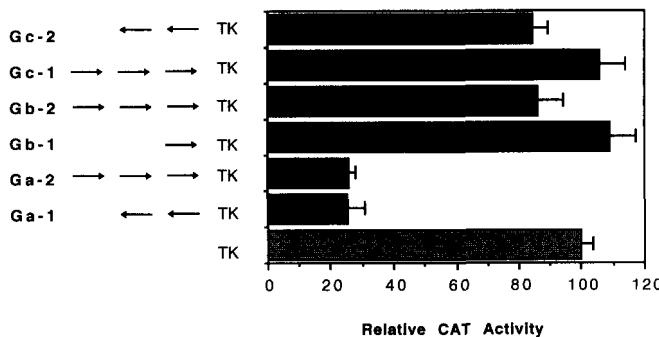


FIG. 5. Transcriptional properties of the proglucagon gene enhancer sequences in InR1-G9 *isl-1*(AS) cells. The proglucagon gene Ga/Gb/Gc enhancer plasmids shown in Fig. 3 were transfected into InR1-G9 *isl-1*(AS) cells (data obtained from transfection of several different InR1-G9 antisense clones were comparable). InR1-G9 transfections were carried out on at least three different occasions (each plasmid in triplicate), and CAT assays were carried out, as described previously (6). The data shown represent the mean \pm S.E. of three separate experiments.

Gb and Gc elements was eliminated in InR1-G9 *isl-1*(AS) cells. Furthermore, the relative levels of proglucagon mRNA transcripts were also lower in the InR1-G9 *isl-1*(AS) cells (data not shown).

DISCUSSION

isl-1 was originally thought to be an insulin gene transcription factor. *In vitro* translated *isl-1* was shown to bind to the TAAT motifs present in the insulin gene FLAT element, and mutation of the TAAT motifs in the insulin gene eliminated the ability of *isl-1* to bind the insulin FLAT element enhancer (15). Nevertheless, although *isl-1* antisera prevented *in vitro* translated *isl-1* from binding to the insulin gene E2 and P1 elements, *isl-1* could not be identified in DNA-protein complexes using nuclear extracts from islet cell lines and the E2 or P1 probes (27). Taken together, these observations suggest that *isl-1* may not form a complex on the insulin gene promoter in cells that actively transcribe the insulin gene *in vivo*. Although *isl-1* also binds sequences in the rat somatostatin promoter (28), recent studies have suggested that the major homeobox protein binding to the somatostatin gene TAAT sites is IDX-1/STF-1 (13, 29).

The subsequent isolation of novel homeobox genes sequences from an insulinoma cDNA library suggests that multiple homeobox transcription factors likely compete for binding to the proglucagon gene promoter AT-rich sequences, since Northern blot analysis demonstrated that several homeobox genes are expressed in InR1-G9 cells and α TC1.9 glucagon-producing islet cell lines (14). Nevertheless, no experiments to date have examined whether islet homeobox genes bind to and regulate the proglucagon gene promoter. The results of our experiments demonstrate that the *isl-1* homeodomain as well as full-length *isl-1* (either translated *in vitro* or present in nuclear extracts) binds to two specific sites in the first 100 bp of the rat proglucagon gene promoter. Deletional and mutational analyses of transfected proglucagon-CAT fusion genes have previously demonstrated that a specific region, designated G1, functions as an islet cell-specific promoter *in vitro* (8). G1 extends from -65 to -100 and contains an inverted TAAT motif on the opposite strand. A second TAAT motif (encompassed by the Ga/Gc sites) is located immediately 3' to G1, approximately 25 bp upstream of the TATA box. Our results demonstrate that the proglucagon gene Ga/Gb/Gc sites are capable of binding the *isl-1* homeodomain *in vitro*. Furthermore, our experiments show that since full-length *in vitro* translated *isl-1* and *isl-1* in islet cell nuclear extracts can also bind to proglucagon gene

sequences, the *isl-1* LIM domains are not necessarily inhibitory for *isl-1* binding to the proglucagon gene promoter *in vitro*.

Analysis of the DNA-binding properties of the TrpE-*isl-1* fusion protein suggests that the *isl-1* binding sites in the proglucagon gene promoter display a higher affinity for *isl-1* than the TAAT sites in the insulin gene FLAT or P1 elements. Competition experiments using excess unlabeled oligonucleotides demonstrated that the E2/P1 sequences were not very effective competitors for *isl-1* binding when using the proglucagon gene Ga/Gb/Gc sequences as probes. In contrast, the proglucagon gene sequences effectively displaced *isl-1* from an E2 probe in EMSA experiments. Taken together, these observations support the hypothesis that the AT-rich sequences in the proximal proglucagon promoter may be target sites for *isl-1* action.

The expression of multiple homeobox genes in a given cell type, taken together with the demonstration that most homeobox genes can bind the core TAAT sequence, makes it difficult to ascertain precisely which homeobox genes exert transcriptionally important effects on a given promoter. The results of our experiments using *isl-1*-depleted antisense InR1-G9 islet cells demonstrate that although the proximal proglucagon gene Ga sequence clearly binds *isl-1* and Ga sequences mediate inhibition of a heterologous promoter, this transcriptional property is not affected by the marked reduction of *isl-1* in the antisense InR1-G9 cells. In contrast, the transcriptional activation mediated by the Gb/Gc sequences was eliminated following transfection of *isl-1* antisense InR1-G9 cell lines, strongly suggesting that *isl-1* mediates the Gb/Gc-dependent activation of the proglucagon gene promoter.

The observation that *isl-1* is detected as part of a high molecular weight complex in InR1-G9 nuclear extracts by EMSA experiments suggests that protein-protein interactions may be important for the formation of a transcriptionally active *isl-1* complex on the proglucagon gene promoter. Increasing evidence implicates the LIM domains as important mediators of protein-protein interaction. Truncation of the LIM domains from the Lmx-1 protein abrogates the synergistic activation of the insulin gene minienhancer in transfected BHK cells (12). The LIM protein RBTN2 and the basic helix-loop-helix protein TAL1 have been shown to form a complex in erythroid cell nuclei (30, 31), and the LIM domains in the human cysteine-rich protein appear to mediate dimerization even in the absence of DNA (32). Mutational analyses have recently shown that the LIM domains in the homeodomain Xlim-1 protein function as negative regulatory elements that constrain Xlim-1 activity (25). Accordingly, it seems reasonable to postulate that the LIM domains of *isl-1* may be important for protein-protein interaction that contributes to complex formation in A cells of the islets. Alternatively, the heterodimer formation observed between the POU protein UNC-86 and the LIM homeobox protein Mec-3 is not dependent on the LIM repeats (33), suggesting that LIM homeobox proteins may contain multiple distinct domains capable of mediating protein-protein interaction *in vitro*. Future experiments should address the identity of the specific partners that complex with *isl-1* that likely contribute to the regulation of proglucagon gene transcription in the endocrine pancreas.

REFERENCES

1. Mojsov, S., Heinrich, G., Wilson, I. B., Ravazzola, M., Orci, L., and Habener, J. F. (1986) *J. Biol. Chem.* **261**, 11880–11889
2. Thorens, B., and Waeber, G. (1993) *Diabetes* **42**, 1219–1225
3. Lee, Y. C., Brubaker, P. L., and Drucker, D. J. (1990) *Endocrinology* **127**, 2217–2222
4. Efrat, S., Teitelman, G., Anwar, M., Ruggiero, D., and Hanahan, D. (1988) *Neuron* **1**, 605–613
5. Lee, Y. C., Asa, S. L., and Drucker, D. J. (1992) *J. Biol. Chem.* **267**, 10705–10708
6. Drucker, D. J., Philippe, J., Jepeal, L., and Habener, J. F. (1987) *J. Biol. Chem.*

- 262**, 15659–15665
7. Gajic, D., and Drucker, D. J. (1993) *Endocrinology* **132**, 1055–1062
 8. Philippe, J., Drucker, D. J., Knepel, W., Jepeal, L., Misulovin, Z., and Habener, J. F. (1988) *Mol. Cell. Biol.* **8**, 4877–4888
 9. Philippe, J., Morel, C., and Prezioso, V. R. (1994) *Mol. Cell. Biol.* **14**, 3514–3523
 10. Knepel, W., Jepeal, L., and Habener, J. F. (1990) *J. Biol. Chem.* **265**, 8725–8735
 11. Knepel, W., Vallejo, M., Chafitz, J. A., and Habener, J. F. (1991) *Mol. Endocrinol.* **5**, 1457–1466
 12. German, M. S., Wang, J., Chadwick, R. B., and Rutter, W. J. (1992) *Genes & Dev.* **6**, 2165–2176
 13. Miller, C. P., McGehee, R. E., Jr., and Habener, J. F. (1994) *EMBO J.* **13**, 1145–1156
 14. Rudnick, A., Ling, T. Y., Odagiri, H., Rutter, W. J., and German, M. S. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 12203–12207
 15. Karlsson, O., Thor, S., Norberg, T., Ohlsson, H., and Edlund, T. (1990) *Nature* **344**, 879–882
 16. Thor, S., Ericson, J., Brännström, T., and Edlund, T. (1991) *Neuron* **7**, 881–889
 17. Dong, J., Asa, S. L., and Drucker, D. J. (1991) *Mol. Endocrinol.* **5**, 1633–1641
 18. Ericson, J., Thor, S., Edlund, T., Jessell, T. M., and Yamada, T. (1992) *Science* **256**, 1555–1560
 19. Wang, M., and Drucker, D. J. (1994) *Endocrinology* **134**, 1416–1422
 20. Schreiber, E., Matthias, P., Muller, M. M., and Schaffner, W. (1989) *Nucl. Acids Res.* **17**, 6419
 21. Varshavsky, A. (1987) *Methods Enzymol.* **151**, 551–565
 22. Takebe, Y., Seiki, M., Fujisawa, J.-I., Hoy, P., Yokota, K., Arai, K.-I., Yoshida, M., and Arai, N. (1988) *Mol. Cell. Biol.* **8**, 466–472
 23. Prost, E., and Moore, D. D. (1986) *Gene (Amst.)* **45**, 107–111
 24. Sánchez-García, I., Osada, H., Forster, A., and Rabbitts, T. H. (1993) *EMBO J.* **12**, 4243–4250
 25. Taira, M., Otani, H., Saint-Jeannet, J.-P., and Dawid, I. B. (1994) *Nature* **372**, 677–679
 26. Sanchez-García, I., and Rabbitts, T. H. (1993) *J. Mol. Biol.* **231**, 945–949
 27. Ohlsson, H., Thor, S., and Edlund, T. (1991) *Mol. Endocrinol.* **5**, 897–904
 28. Leonard, J., Serup, P., Gonzalez, G., Edlund, T., and Montminy, M. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 6247–6251
 29. Leonard, J., Peers, B., Johnson, T., Ferri, K., Lee, S., and Montminy, M. R. (1993) *Mol. Endocrinol.* **7**, 1275–1283
 30. Valge-Archer, V. E., Osada, H., Warren, A. J., Forster, A., Li, J., Baer, R., and Rabbitts, T. H. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 8617–8621
 31. Wadman, I., Li, J., Bash, R. O., Forster, A., Osada, H., Rabbitts, T. H., and Baer, R. (1994) *EMBO J.* **13**, 4831–4839
 32. Feuerstein, R., Wang, X., Song, D., Cooke, N. E., and Liebhäber, S. A. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 10655–10659
 33. Xue, D., Tu, Y., and Chalfie, M. (1993) *Science* **261**, 1324–1328