Synergistic activation of the insulin gene by a LIM–homeo domain protein and a basic helix–loop–helix protein: building a functional insulin minienhancer complex

Michael S. German, Juehu Wang, Robert B. Chadwick, and William J. Rutter
Hormone Research Institute, University of California, San Francisco, California 94143-0534 USA

The distal portion of the rat insulin I gene 5′-flanking DNA contains two sequence elements, the Far and FLAT elements, that can function in combination, but not separately, as a β-cell-specific transcriptional enhancer. We have isolated several cDNAs encoding proteins that bind to the FLAT element. Two of these cDNAs, cdx-3 and lmx-1, represent homeo box containing mRNAs with restricted patterns of expression. The protein encoded by lmx-1 also contains two amino-terminal cysteine/histidine-rich “LIM” domains. Both cdx-3 and lmx-1 can activate transcription of a Far/FLAT-linked gene when expressed in a normally non-insulin-producing fibroblast cell line. Furthermore, in fibroblasts expressing transfectected β-cell lmx-1, the addition of the Far-binding, basic helix–loop–helix protein shPan-1 (the hamster equivalent of human E47) causes a dramatic synergistic activation. ShPan-1 causes no activation in fibroblasts expressing transfected cdx-3 or the related LIM-homeodomain protein isl-1. Deletion of one or both of the LIM domains from the 5′ end of the lmx-1 cDNA removes this synergistic interaction with shPan-1 without any loss of basal transcriptional activation. We conclude that β-cell lmx-1 functions by binding to the FLAT element and interacting through the LIM-containing amino terminus with shPan-1 bound at the Far element. These proteins form the minimal components for a functional minienhancer complex.

[Key Words: Insulin gene; cdx; lmx-1; islet amyloid polypeptide gene; transcription; LIM domain; homeo box]

Received June 22, 1992; revised version accepted August 13, 1992.

In adult mammals the insulin gene is transcribed exclusively in the β-cells of the pancreatic islets of Langerhans (Clark and Steiner 1969; Giddings et al. 1985). This unique expression results in part from the cell-specific nature of the insulin gene 5′-flanking DNA or promoter, which limits the expression of a linked gene to the β-cell (Walker et al. 1983; Hanahan 1985). Deletion and substitution analyses of the insulin promoter have revealed multiple cis-acting sequence elements that regulate transcriptional activity (Edlund et al. 1985; Karlsson et al. 1987; Crowe and Tsai 1989; Whelan et al. 1989; Boam et al. 1990). These sequence elements presumably function by serving as recognition sites for sequence-specific DNA-binding proteins that, in turn, interact with the transcriptional machinery to activate transcription. The β-cell nucleus contains both ubiquitous and apparently unique protein complexes that bind the cis-acting elements in the insulin promoter (Moss et al. 1988; Ohlsson et al. 1988; Boam et al. 1990; Whelan et al. 1990; Aronheim et al. 1991; German et al. 1992). The cell specificity of the insulin promoter presumably results from the unique combination of nuclear complexes formed by these protein–DNA interactions.

A short portion of the rat insulin I promoter (the FF minienhancer, see Fig. 1) between −247 and −197 bp upstream from the transcription start site can act as a transcriptional activator when linked to a heterologous promoter (Karlsson et al. 1989); it responds to glucose similarly to the intact promoter (German et al. 1990); and it contains at least two sequence elements, the Far element (−239 to −228) and the FLAT element (−222 to −208), that are sensitive to mutation in the intact promoter (Karlsson et al. 1987) or the minienhancer (German et al. 1992). The Far and FLAT sequence elements function synergistically: Neither the Far element nor the FLAT element can function alone, but both are required for minienhancer function. In addition, the FLAT element itself is composed of two functionally distinct elements, FLAT-E and FLAT-F, either one of which is sufficient for full minienhancer function (German et al. 1992). The Far element binds an apparently endocrine cell-restricted nuclear complex that also binds the similar, proximal Nir element, which is also conserved in the rat insulin II and human insulin promoters (Moss et al. 1988; Ohlsson et al. 1988; Boam et al. 1990, Aronheim et al. 1991). The basic helix–loop–helix proteins

In adult mammals the insulin gene is transcribed exclusively in the β-cells of the pancreatic islets of Langerhans (Clark and Steiner 1969; Giddings et al. 1985). This unique expression results in part from the cell-specific nature of the insulin gene 5′-flanking DNA or promoter, which limits the expression of a linked gene to the β-cell (Walker et al. 1983; Hanahan 1985). Deletion and substitution analyses of the insulin promoter have revealed multiple cis-acting sequence elements that regulate transcriptional activity (Edlund et al. 1985; Karlsson et al. 1987; Crowe and Tsai 1989; Whelan et al. 1989; Boam et al. 1990). These sequence elements presumably function by serving as recognition sites for sequence-specific DNA-binding proteins that, in turn, interact with the transcriptional machinery to activate transcription. The β-cell nucleus contains both ubiquitous and apparently unique protein complexes that bind the cis-acting elements in the insulin promoter (Moss et al. 1988; Ohlsson et al. 1988; Boam et al. 1990; Whelan et al. 1990; Aronheim et al. 1991; German et al. 1992). The cell specificity of the insulin promoter presumably results from the unique combination of nuclear complexes formed by these protein–DNA interactions.
German et al.

Pan-1 and Pan-2 [equivalent to human E47 and E12 (Murre et al. 1989)] bind the Far element [Nelson et al. 1990; Walker et al. 1990], and antisera to Pan-1 and Pan-2 recognize the endocrine-specific Far-binding nuclear complex [Aronheim et al. 1991; Cordle et al. 1991; German et al. 1991; Sheih and Tsai 1991].

The FLAT element binds several nuclear complexes, at least two of which have restricted tissue distribution [Aronheim et al. 1991; German et al. 1992]. The FLAT-E, but not the FLAT-F, element is well conserved in all known insulin promoters [Steiner et al. 1985], and similar elements are found in the β-cell-specific human islet amyloid polypeptide [IAPP] promoter [Nishi et al. 1989]. The IAPP promoter FLAT-like sequence can bind all of the nuclear complexes that bind the FLAT element of the rat insulin I promoter [German et al. 1992]. The selectively expressed cDNA isl-1 encodes a LIM–homeo domain protein capable of binding the FLAT element [Karlsson et al. 1990], but antisera to isl-1 does not recognize any of the major nuclear FLAT-binding complexes [Aronheim et al. 1991]. In this paper we report the characterization of cDNAs encoding other FLAT-binding proteins and study the interactions of these proteins with the FF minienhancer and Far-binding proteins.

Results

Isolation and sequencing of cdx-3 and lmx-1

We used a multimerized 32P-labeled DNA probe containing the FLAT element (see Fig. 1) to screen 500,000 plaques from an unamplified λgt11 cDNA library prepared from the Syrian hamster insulinoma cell line HIT T-15 M2.2.2. Twelve independent phage clones encoding proteins that specifically bound the FLAT element probe were identified. Three of these clones, λF4, λF11, and λF12, are described here. DNA sequence revealed that the 2.1-kb λF12 cDNA insert was identical to the λF4 insert except for an additional 102 bp of 5′ sequence (Fig. 2). The first potential ATG initiator codon in both cDNAs (at +8 bp in λF4) would initiate a 13-amino-acid peptide. A second ATG (at +40 bp in λF4) is in-frame with the λgt11 lacZ translation frame and encodes a 313-amino-acid protein with a predicted molecular mass of 34 kD. Four additional cDNA clones obtained by hybridization screening of the HIT λgt11 library with the λF4 cDNA also contained both ATGs (Fig. 2). Using antisera directed at the carboxyl terminus, we tested for the presence of the 34-kD translation product in HIT nuclear extract and found a DNA-binding protein with electrophoretic mobility identical to the in vitro-produced 34-kD F4 protein (data not shown). For the remainder of this discussion, we will refer to the 34-kD protein as cdx-3 because of its similarity to the other cdx genes.

The amino acids from 185 to 245 of cdx-3 form the homeo domain. The F4 homeo domain is closely related to the homeo domains found in the Drosophila gene caudal (cad) (Mlodzik et al. 1985), the mouse genes cdx-1 [for caudal-type homeo box] (Duprey et al. 1988) and cdx-2 [James and Kazenwadel 1991], the chicken gene CHox-cad (Frumkin et al. 1991), and the Caenorhabditis elegans gene ceh-3 (Burglin et al. 1989) (See Fig. 4A, be-

![Figure 1](https://example.com/figure1.png)

**Figure 1.** Insulin and islet amyloid polypeptide minienhancer sequences. (A) The position of the FF minienhancer within the rat insulin I gene 5′-flanking DNA is shown. The shaded sequence was used to construct the multimerized probe for screening for FLAT-binding clones. (B) The identical nucleotides in the related minienhancers from the rat insulin I and II genes are boxed. (C) The sequences of the wild-type and mutant insulin and islet amyloid polypeptide minienhancers are shown. Changes in the mutant minienhancers are underlined.
low). Interestingly, all of these genes are expressed in the gut: \textit{cdx-1} and \textit{cdx-2} in the intestinal epithelium of the developing and adult mouse, \textit{cad} in the posterior midgut, hindgut, and Malpighian tubes of \textit{C. elegans} (Mlodzik and Gehrig 1987), and \textit{CHox-cad} in the epithelial lining of the early developing chicken gut and yolk sac. Outside of the homeo domain region, \textit{cdx-3} shares some sequence similarity with the other \textit{cad}-related genes at the amino terminus but otherwise diverges from these genes (for comparisons, see Frumkin et al. 1991).

The DNA sequence of the \textit{\lambda}F11 cDNA insert did not reveal an obvious in-frame initiator ATG codon [Fig. 3]. The F11 cDNA was therefore used to screen the \textit{Kgt1} library by hybridization. A recombinant \lambda phage with a larger, 3.8-kb cDNA insert was obtained (\textit{\lambda}F11.11.6; Fig. 3). This larger cDNA has a large open reading frame starting with a potential initiator ATG at 117 bp from the 5' end and encodes a 382-amino-acid protein with a predicted molecular mass of 43 kDa.

The amino-terminal half of the predicted F11 [or \textit{lmx-1} for LIM–homeo box] amino acid sequence contains two tandem copies of a cysteine/histidine-rich "LIM" domain. Several other homeo box genes also contain LIM domains: the \textit{C. elegans} genes \textit{me-3} (Way and Chalfie 1988) and \textit{lin-11} (Freyd et al. 1990), the rat gene \textit{isl-1} (Karlsson et al. 1990), the \textit{Xenopus laevis} gene \textit{Xlim-1} (Taira et al. 1992), and the \textit{Drosophila} gene \textit{apterous} (\textit{ap}) (Cohen et al. 1992) [see Fig. 4B]. The \textit{lin-11} LIM domain has been shown to bind iron-sulfur and zinc (Li et al. 1991). The \textit{lmx-1} homeo domain shares certain features of the other LIM protein homeo domains but is not closely related to any of these homeo domains [Fig. 4A].

In the process of screening for additional \textit{cdx-3} and \textit{lmx-1} cDNA clones, we obtained a large number of weakly hybridizing signals. Some of these related cDNA clones encode proteins capable of binding the FLAT element (data not shown). Included among these is the hamster equivalent of the rat \textit{isl-1} cDNA [Karlsson et al. 1990]. These clones, in addition to the clones obtained in the original screen, offer further evidence of the complexity of protein–DNA interactions at the FLAT element in \textit{HIT} cells.

\textit{cdx-1} and \textit{lmx-1} demonstrate similar DNA-binding preferences

The coding portions of the \textit{cdx-3} and \textit{lmx-1} cDNAs were inserted into T7-driven transcription vectors, and transcribed and translated in vitro. The resulting protein products were tested for DNA binding by electrophoretic mobility shift assay (EMSA; see Fig. 5, below) using a labeled DNA probe containing the Far–FLAT region of the rat insulin I gene [FF minienhancer; Fig. 1]. Although in vitro-produced \textit{cdx-3} and \textit{lmx-1} bind the probe with high affinity, \textit{isl-1} only binds significantly when the concentration of the nonspecific competitor poly[dI-C] is reduced 100-fold [Fig. 5A]. Analysis with mutant-binding sites demonstrates that both \textit{cdx-3} and \textit{lmx-1} can bind at either the FLAT-E or the FLAT-F site, although both proteins bind best at the FLAT-E site [Fig. 5C]. When a mixture of \textit{cdx-3} and \textit{lmx-1} is bound to the FF probe, no significant new complexes are seen [data not shown].

Both proteins bind to the FLAT-related sites (for se-
sequences, see Fig. 1) in the rat insulin II and hAPP genes (Fig. 5B). Lmx-1 produces a prominent band with decreased electrophoretic mobility when bound to the hAPP probe. This may reflect additional lmx-1 proteins binding to the multiple FLAT-related sites present within this sequence. These experiments were performed with roughly equivalent amounts of protein, as judged by [35S]methionine labeling.

cdx-3 and lmx-1 have restricted tissue distribution

By Northern analysis, the cdx-3 cDNA hybridizes with a 2.1-kb mRNA in several endocrine lines but not in the hamster fibroblast cell lines BHK21 and CHO (Fig. 6A). The multiple minor bands seen here are not present when a shorter cdx-3 probe missing the distal homeo domain, glutamine-rich domain, and carboxy-terminal tail is used (data not shown). No signal was seen when poly(A)+ RNAs from several tissues including pancreas were probed by Northern analysis with the cdx-3 cDNA (data not shown).

We were unable to detect lmx-1 mRNA in non-insulin-producing cell lines by Northern analysis, and the signals from HIT-T15 and RIN.M5F cell poly(A)+ RNAs were very weak as well (data not shown). By RNase protection, lmx-1 mRNA is seen in HIT-T15 cells but not in BHK21 or any hamster tissues including pancreas (Fig. 6B). No signal was seen when a shorter tail is used (data not shown). No signal was seen when a shorter tail is used (data not shown). No signal was seen when a shorter tail is used (data not shown). No signal was seen when a shorter tail is used (data not shown).

Figure 3. lmx-1 DNA sequence. The nucleotide and derived amino acid sequences of the A/F11.11.6 cDNA insert are shown without 2.6 kb of additional 3'-untranslated sequence. Arrows indicate the 5' and 3' ends of A/F11, the 5' ends of lmx-1, and the 3' ends of lmx-1.D2 and lmx-1.D3, and the 3' ends of lmx-1.D4 and lmx-1.D5. A broken line overlies the homeo domain. The two LIM domains are boxed separately.

Figure 4. Homeo domain protein comparisons. (A) The cdx-3 homeo domain sequence is compared with other members of the cad class of homeodomains, and the lmx-1 homeo domain sequence is compared with other members of the LIM class of homeodomains. (B) The LIM domains from the LIM-homeo class of homeodomains, and the cdx-3 domain proteins are compared.
Insulin minienhancer-binding proteins

Figure 5. DNA binding of cdx-3, lmx-1, and shisl-1. In vitro-produced proteins were tested by EMSA for the ability to bind the labeled DNA probes shown. The control protein is the product of transcription and subsequent translation of the transcription vector without any cDNA insert. (A) Binding to the FF minienhancer was tested in the presence of 200 or 2 μg/ml of the nonspecific competitor poly[d(I-C)]/[d(I-C)]. (B,C) Binding to the different minienhancer probes is compared using equal probe concentrations and equal volumes of the protein translation products in all lanes. Roughly equal efficiency of protein translation was demonstrated by [35S]methionine labeling.

6C). A shPan-1 probe was used to test these samples for relative mRNA quality and quantity. A pattern similar to previous Northern analyses was seen (Nelson et al. 1990, German et al. 1991): The two cell lines HIT-T15 and BHK21 had equal shPan-1 mRNA levels; the tissue samples had lower, but roughly equal, levels (data not shown).

cdx-3 expression in selected hamster tissues was also tested by RNase protection and revealed the presence of cdx-3 mRNA in the intestine. A faint band is protected by the cdx-3 probe in testes mRNA as well (Fig. 6B). Because islets represent only a small percentage of the adult pancreas, it is unlikely that modest levels of expression of lmx-1 or cdx-3 in islets could be detected by RNase protection analysis using 257-bp 32p-labeled probe that protects 179 bp of the cdx-3 mRNA. Five micrograms of poly[4] RNA was added from each tissue shown. (C) RNase protection analysis was performed using a 234-bp 32p-labeled probe that protects 156 bp of the lmx-1 mRNA. Five micrograms of poly[4] RNA was added from each tissue shown.

Figure 6. Selective expression of the cdx-3 and lmx-1 mRNAs. (A) By Northern blot analysis, 15 μg of poly[4] RNA from the tumor cell lines shown was annealed to a full-length cdx-3 cDNA probe. (B) RNase protection analysis was performed using a 257-bp 32p-labeled probe that protects 179 bp of the cdx-3 mRNA. Five micrograms of poly[4] RNA was added from each tissue shown. (C) RNase protection analysis was performed using a 234-bp 32p-labeled probe that protects 156 bp of the lmx-1 mRNA. Five micrograms of poly[4] RNA was added from each tissue shown.
German et al.

this means. This was confirmed with antisera directed at the carboxyl termini of cdx-3 and lmx-1; both antisera recognize antigens present in the islets but not in the exocrine cells of adult mouse pancreases (G. Gittes, M.S. German, and W.J. Rutter, unpubl.).

cdx-3 and lmx-1 activate portions of the insulin and IAPP enhancers

The coding portions of the cdx-3, lmx-1, and isl-1 cDNAs were inserted into expression plasmids driven by the human cytomegalovirus [CMV] immediate early gene promoter. This allows expression of the full-length proteins in eukaryotic cells. These expression plasmids were transfected into the non-insulin-producing hamster fibroblast cell line BHK21, along with a reporter plasmid expressing the firefly luciferase gene. A prolactin minimal promoter lies immediately upstream of the luciferase gene in the reporter plasmids. Five copies of the wild type or mutant rat insulin I FF minienhancer, the rat insulin II R2 minienhancer, or the human islet amyloid polypeptide AX minienhancer were inserted upstream of the prolactin promoter [for sequences, see Fig. 1]. This permitted us to test the ability of these proteins to activate transcription by the minienhancers.

Normal BHK21 cells do not activate the FF minienhancer [German et al. 1992] (Fig. 7A). The presence of cotransfected lmx-1, however, markedly activates minienhancer-directed expression of the luciferase reporter function. Cotransfected cdx-3 also activates the FF construct, although cdx-3 activates the enhancerless prolactin luciferase [prl–Luc] construct to a lesser degree as well. Cotransfected isl-1 does not activate luciferase expression. Similar results are seen when the prolactin promoter is replaced with the herpes simplex virus I thymidine kinase promoter. Cotransfection of lmx-1 and cdx-3 together results in levels of activation intermediate between the levels seen with either cDNA alone [data not shown].

Full lmx-1 activation requires an intact FLAT-E site, and the minienhancer with both sites mutated [EF] is not activated by lmx-1 [Fig. 7B]. lmx-1 also activates the related hIAPP minienhancer [see sequence, see Fig. 1] which has three sequence elements similar to FLAT-E and FLAT-F. lmx-1 does not activate the R2 minienhancer, which has a conserved FLAT-E site but no FLAT-F or Far element. Interestingly, cdx-3 does not activate the EF or R2 constructs despite the fact that it activates the enhancerless construct. This suggests that the activation of the enhancerless prl–Luc construct may involve flanking sequences uniquely juxtaposed in that construct and that the EF and R2 constructs are better negative controls.

lmx-1 and shPan synergistically activate the insulin enhancer/promoter

The shPan-1 cDNA, which encodes a Far element-binding protein identical to the human E47 protein, does not activate the FF minienhancer by itself [Fig. 8A]. In the

Figure 7. Activation of the minienhancer by lmx-1 and cdx-3 in BHK21 cells. A plasmid expressing the luciferase gene under control of a minimal prolactin promoter and five copies of the minienhancer was cotransfected with a plasmid expressing the cDNA, and luciferase activity was measured in cell extracts 48 hr later. Data points [labeled control] were cotransfected with an expression plasmid without a cDNA insert. Luciferase activity in the cells transfected with the enhancerless prl–Luc plasmid and the control expression plasmid [the first data point in A] was arbitrarily set at 1.0.
Insulin minienhancer-binding proteins

Figure 8. Synergistic activation of the minienhancer and intact insulin promoter by lmx-1 and Pan-1 in BHK21 cells. (A) A plasmid expressing the luciferase gene under the control of a minimal prolactin promoter and five copies of the minienhancer was cotransfected with a plasmid expressing the cDNAs and an additional expression plasmid with either a Pan-1 cDNA insert (solid bars) or no cDNA insert (hatched bars). Luciferase activity in the cells transfected with the enhancerless prl-Luc plasmid and the control expression plasmid (the first data point) was arbitrarily set at 1.0. (B) A plasmid expressing the luciferase gene under the control of 410 bp of rat insulin I gene 5′ flanking DNA was cotransfected with plasmids expressing the cDNAs shown. Luciferase activity in the cells transfected with the control expression plasmid (the first data point) was arbitrarily set at 1.0.

presence of cotransfected lmx-1, however, the addition of shPan-1 results in a dramatic further activation of the minienhancer. Minienhancers mutated in either the Far element or the entire FLAT element (mutation EF) do not display this synergistic effect. Significant synergistic activation is not seen when shPan-1 is cotransfected with either cdx-1 or isl-1 [Fig. 8A]. The shPan-2 cDNA [a splice variant of shPan-1 equivalent to human E12] gave quantitatively similar results [data not shown].

When the intact rat insulin I enhancer/promoter (−410 to +1 bp relative to the transcription start site) is inserted in place of the minienhancer–prolactin promoter in the luciferase reporter plasmid, shPan-1 alone cannot activate luciferase expression. Cotransfected lmx-1, however, activates the intact enhancer/promoter, and shPan-1 plus lmx-1 causes further activation [Fig. 8B]. To a lesser degree, cdx-3 also activates the intact enhancer/promoter [data not shown].

Deletion analysis of lmx-1 and shPan

Expression of a shPan-1 protein with a large amino-terminal deletion, shPan.SH, does not synergistically activate the FF minienhancer in the presence of lmx-1 [Fig. 9A]. The shPan.SH protein, which has 508 amino-terminal amino acids removed, still avidly binds the Far element when tested in vitro [M. Blanar and W. Rutter, unpubl.]. A carboxy-terminally deleted shPan cDNA, shPan.SS, cannot bind DNA [German et al. 1991] and also cannot synergistically activate the FF minienhancer [Fig. 9A].

We made several deletions from both ends of the lmx-1 cDNA. The amino-terminal deletions D2 and D3 remove the first and second LIM domains, respectively [Fig. 3]. Deletion D4 removes the carboxy-terminal 110 amino acids after the glutamine-rich domain. Deletion D5 removes part of the homeo domain, the glutamine stretch, and the carboxyl terminus. All of these truncated proteins were produced in vitro, and all but lmx-1.D5 bind as well as the full-length protein to the FF minienhancer probe. The lmx-1.D5 protein does not bind the minienhancer [data not shown].

When expressed in BHK21 cells, the two amino-terminally truncated proteins lmx-1.D2 and lmx-1.D3 activate the FF minienhancer, but neither lmx-1.D2 nor lmx-1.D3 synergistically activates the FF minienhancer in the presence of shPan-1. Neither of the carboxy-terminally truncated proteins can activate the FF minienhancer; however, lmx-1.D4, but not lmx-1.D5, still gives synergistic activation in the presence of shPan-1 [Fig. 9B]. These results suggest that the amino end of lmx-1 with its LIM domains is necessary for the synergistic activation, whereas the carboxy end is required for activation in the absence of cotransfected Pan.
We also replaced the entire amino-terminal portion of \( lmx-1 \) with the equivalent portions of the \( isl-1 \) cDNA, such that a chimeric protein is produced with the \( isl-1 \) LIM domains linked to the F11 homeo domain and carboxyl terminus. This chimeric construct functions like \( lmx-1.D2 \); it gives basal but not synergistic activation of the FF minienhancer [Fig. 9B].

When tested in vitro by EMSA, \( lmx-1 \) and shPan-1 do not bind DNA synergistically: Neither protein binds at a higher affinity or produces different complexes in the presence of the other protein. These binding experiments were performed with mixed proteins and with cotranslated proteins in several different binding conditions. We were also unable to demonstrate any direct contact between \( lmx-1 \) and shPan-1 by immunoprecipitation [data not shown]. These results are not surprising because antisera to the Pan proteins do not recognize a complex in insulinoma nuclear extract that is appropriate in size for \( Pan + lmx-1 \) [Aronheim et al. 1991; Cordle et al. 1991; German et al. 1991, Sheih and Tsai 1991]. The presence of high levels of the two LIM domains without a homeo domain [i.e., \( lmx-1.D5 \)] does not interfere with the DNA binding of full-length \( lmx-1 \) or shPan-1 in vitro or with synergistic activation by full-length \( lmx-1 \) and shPan-1 in transfection experiments in BHK21 cells [data not shown].

**Discussion**

**Components of the minienhancer complexes**

The present studies reinforce the concept that multiple proteins bind and interact at the insulin minienhancer.
The ability to clone these proteins provides potential components of the minienhancer machinery, but DNA-binding characteristics alone do not distinguish the relative roles of these proteins. Several lines of evidence suggest that isl-1, the first FLAT element-binding factor described, may play a limited role in the function of the Far-FLAT minienhancer. The in-vitro-produced isl-1 protein has a low specific affinity for the FLAT element; transfected isl-1 does not activate the minienhancer; and neither the intact isl-1 nor its LIM domains linked to the lmx-1 homeo domain synergistically interact with Pan-1. These data, along with evidence that isl-1 has higher affinity for other binding sites (Aronheim et al. 1991) and is expressed in a number of different cell types (Dong et al. 1991; Thor et al. 1991), suggest that isl-1 may play a more important role elsewhere. However, we cannot exclude the possibility that by interacting with some other protein partner isl-1 could significantly activate the minienhancer.

The relative roles of cdx-3 and lmx-1 in the minienhancer complex are difficult to judge based on the transfection data alone. Our data do not support the concept that cdx-3 and lmx-1 are partners in a single large complex. Because they can form independent complexes, the two genes may display distinct temporal patterns and relative levels of expression during development and therefore may play different roles at different stages of development. lmx-1 displays the characteristics of a major activator of the minienhancer and thus of insulin gene expression. It has a high affinity for the FLAT-binding sites, it dramatically activates the minienhancer in conjunction with Pan, and it appears to have a tightly restricted pattern of expression.

cdx-3 also binds the FLAT sites with high affinity and has a restricted pattern of expression, but it does not synergistically interact with Pan. If both proteins are expressed simultaneously, cdx-3 may modulate the Pan/lmx-1 interaction by competing with lmx-1 for the FLAT-binding sites. cdx-3 may also have an as yet unidentified activation partner analogous to Pan.

The intestinal expression of the c-ad-related genes and the presence of the related cdx-3 protein in islet cells and intestine is intriguing. A unique gut–islet relationship is suggested by the selective expression in both tissues of several polypeptide hormones, including IAPP, glucagon, somatostatin, and pancreatic polypeptide (Green et al. 1989). The coexpression of insulin and gut peptide hormones in islet cell tumors is common (Wynick et al. 1988; Green et al. 1989). During development, the mammalian pancreas forms initially as a pouch or bud of intestinal epithelium that eventually differentiates into the islets of Langerhans and the exocrine acini that form the adult pancreas (Wessells and Cohen 1967; Pictet et al. 1972). The ability of the islet endocrine cells to transcribe the gastrointestinal endocrine cell genes suggests that some common genetic or developmental characteristics persist in these differentiated cells. It will be interesting to see where and when the different cdx genes are expressed during the development of the gut and pancreas.

The large number of FLAT-binding complexes in HIT nuclei and the additional FLAT-binding clones suggests the existence of other functionally important components of the FLAT-binding complexes. Furthermore, there is evidence for additional Far-binding proteins (Aronheim et al. 1991; Cordle et al. 1991; German et al. 1991; Sheih and Tsai 1991), raising the possibility of numerous protein–protein and protein–DNA interactions.

**Functional interaction between lmx-1 and the Pan proteins**

Multicellular organisms regionally express numerous homeo domain proteins (for review, see Scott et al. 1989), presumably resulting in activation of different promoters in different regions. Because homeo domain proteins have broad and overlapping DNA-binding preferences, differential promoter function must not result from exclusive use of different DNA-binding sites. The presence of a homeo domain-binding site and a selectively expressed homeo domain protein cannot alone explain cell-specific function of a promoter because other cell types contain homeo domain proteins that bind the same site. Subtle differences in DNA-binding preferences of selectively expressed homeo domain proteins might change the overall phenotype of a cell, but this is not likely to explain the all-or-none pattern of expression of cell-specific genes such as insulin. The functional specificity of most of these proteins must therefore lie outside the DNA-binding domain.

One means of increasing the promoter- or cell-specific function of these proteins is through specific interactions with other proteins binding to adjacent sites. This appears to be how lmx-1 specifically activates the FF minienhancer and the insulin promoter. We have demonstrated that synergistic activation of the minienhancer by lmx-1 and Pan-1 (1) requires DNA binding by both proteins to their adjacent sites, (2) requires the non-DNA-binding amino end of Pan-1 and the lmx-1 LIM domains, (3) is not blocked by overexpression of the LIM domains alone, and (4) does not result from synergistic DNA binding, at least as assayed in vitro. Although we were unable to demonstrate direct physical contact between lmx-1 and Pan-1, such an interaction may exist but may not withstand the harsh means we used to test for it or may require the presence of other cellular components or protein modifications not provided by the in vitro-translated proteins. Alternatively, the lmx-1/Pan-1 interaction may be indirect. One simple explanation could be a linking protein that interacts simultaneously with lmx-1 and Pan-1.

Although we cannot exclude the possibility that the small peptide sequence on the amino terminus of lmx-1 may be required or that the isl--lmx fusion protein may be improperly folded or processed, the requirement for both intact LIM domains and the inability of the isl-1 LIM domains to substitute for the lmx-1 LIM domains suggests that the Pan/lmx-1 interaction is specific for the lmx-1 LIM domains. As a result, specific function of the minienhancer results from this complex interac-
tion—governed by the linked binding sites and the presence of the required proteins—and not from the expression of the DNA-binding proteins alone.

BH21 cells expressing the Pan-1 and Imx-1 cDNAs display a new feature seen previously only in insulin-producing cells. The tripartite Far–FLAT sequence now functions as an active enhancer; and, just as in HIT cells, it requires an intact Far element and at least one copy of the FLAT-E or FLAT-F sites for full function. Clearly Imx-1, alone or with Pan-1, can switch the insulin promoter from an inactive to an active state in these cells. Native Imx-1 likely plays this same role in β-cells along with other, possibly related proteins. Although we may not have identified all of the protein components of the insulin minienhancer machinery, the Pan proteins and the cdx-3 and Imx-1 proteins are minimal sufficient components for insulin minienhancer function. The interaction between the ubiquitous Pan proteins and the β-cell Imx-1 protein is further evidence of the tightly restricted interactions necessary for construction of a functioning minienhancer complex.

Materials and methods

Isolation of cDNA clones

Construction of the Agt11 HIT T15 M2.2.2 cDNA library and the method used to screen for DNA-binding clones have been described previously [German et al. 1991]. The same filters used previously to screen for Ntr element-binding clones [shPan-1; German et al. 1991] were stripped and rescreened with the multimerized FLAT probe shown in Figure 1. Sequence of both strands of each cDNA was obtained by the Sanger dideoxy sequencing method.

Cell lines and transfections

The Syrian hamster kidney fibroblast line BH21 [Karlsson et al. 1987] was grown in DMEM16 medium with 3 grams of glucose per liter to a density of 1 x 10⁶ cells per 100-mm plate before transfection. Transfection was performed by the calcium-phosphate technique [Karlsson et al. 1987] using 10 μg of double cesium chloride-purified luciferase plasmid DNA and 1 μg of each cotransfected expression plasmid. Cells were harvested, and protein extracts were prepared 48 hr after transfection. Luciferase assays were performed with 100 μg of protein extract [deWet et al. 1987]. Each data point represents the mean of at least three independent transactions ± S.E.

αTC1 and βTC1, transgenic mouse α- and β-cell lines, respectively, were obtained from D. Hanahan [University of California at San Francisco]. CHO, a Chinese hamster ovarian fibroblast line; PC12, a rat adrenal medullary cell line; and RIN.MSF, a rat β-cell line, were obtained from the University of California Cell Culture Facility.

Plasmid constructions

The luciferase plasmids are based on the plasmid pMG5, which contains the firefly luciferase cDNA from +30 to +1747 bp [kindly provided by S. Subramani, University of California at San Diego; deWet et al. 1987], followed by an SV40 polyadenylation signal in pUC19. Upstream of the HindIII restriction endonuclease recognition site on the 5’ end of the luciferase cDNA is a polylinker with multiple restriction endonuclease recognition sites preceded by two additional SV40 polyadenylation signals. The construction of the minienhancer multimers has been described previously [German et al. 1992]. These five-copy multimers were inserted in a coding orientation upstream of a minimal rat prolactin promoter [–36 to +34 bp relative to the transcription start site of the rat prolactin gene [Nelson et al. 1988]] linked to the 5’ end of the luciferase cDNA in pMG5. The –410 insulin promoter plasmid was constructed similarly by inserting the –410 to +1-bp rat insulin I genomic DNA fragment upstream of the luciferase cDNA.

The transcription and CMV expression vectors are based on two parallel plasmids, pBAT9 and pBAT7, respectively. The pBAT9 plasmid was derived from pBlueScript SK + [Stratagene Cloning Systems, LaJolla, CA] and has an SV40 polyadenylation signal inserted into the SacI site and a consensus translation start site and initiator ATG codon with an in-frame EcoRI restriction endonuclease recognition site as described previously for the plasmid pTR1 [German et al. 1991]. The cDNAs were inserted such that translation would start with the predicted initiator ATG codons shown in Figures 2 and 3 using the consensus translation start site. The pBAT7 plasmid is identical to pBAT9, with the addition of the CMV immediate early gene promoter [Boshart et al. 1985] upstream of the ATG. The shPan-1 pBAT7 construct differs from the other pBAT7-based plasmids because it includes 44 bp of the 5’-untranslated sequence of shPan-1 and therefore does not have the ideal consensus translation start signal. The RNase protection plasmids were constructed by inserting a 179-bp PsI–HindIII fragment from cdx-3 or a 156-bp PsI–HindIII fragment from Imx-1 into the appropriate restriction endonuclease recognition sites in the pBlueScript KS + polylinker [Stratagene Cloning Systems, LaJolla, CA].

The shPan-1.SH deletion that removes helix 2 from shPan-1 [shPan-1.SS] has been described previously [German et al. 1991]. The shPan-1.SH deletion was kindly provided by M. Blanar [University of California at San Francisco] and inserted into the EcoRI site in pBAT7. The 5’ Imx-1 deletions, D2 and D3, were constructed with a single ATG codon replacing the deleted cDNA sequence and were inserted into pBAT9 and pBAT7. The 3’ deletions, D4 and D5, were constructed by replacing the deleted cDNA sequence with an in-frame stop codon and were inserted into pBAT9 and pBAT7. The shFS-1.C deletion was created by replacing the 5’ end of the Imx-1 cDNA up to the codon for cysteine-143 with the 5’ end of the shFS-1 cDNA up to the codon for cysteine-129 [see Karlsson et al. 1990] and inserted into pBAT7.

Northern analysis

The cell lines shown in Figure 7 were grown to ~75% confluency, and RNA was isolated with the Fast Track mRNA isolation kit [Invitrogen Corp., San Diego, CA]. The RNA was heat denatured, separated on a formaldehyde-polyacrylamide gel, and blotted onto Genescreen nylon membrane [New England Nuclear Research Products, Boston, MA]. The membrane was probed at high stringency with random primer 32P-labeled cdx-3 cDNA, as directed by the membrane manufacturer.

RNase protection

RNA was isolated using RNAzol (Tel-Test Inc., Friendswood, TX), followed by oligo(dT)-Sepharose poly[A]+ RNA purification. Single-stranded complimentary RNase protection probes were produced by transcription with T7 RNA polymerase of the linearized pBlueScript plasmid templates described above. The
Acknowledgments

We thank Julia Varley and Leslie Spector for preparation of the manuscript, Olga Venekie for synthesis of oligonucleotides, Giulia Kennedy for critical reading of the manuscript, and J. Edman and C. Nelson for helpful discussions.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

Note added in proof

Sequence data described in this paper have been submitted to the EMBL/GenBank data libraries.

References

Aronheim, A., H. Ohlsson, C.W. Park, T. Edlund, and M.D. Walker. 1991. Distribution and characterization of helix-loop-helix enhancer-binding proteins from pancreatic β-cells and lymphocytes. Nucleic Acids Res. 19: 3893–3899.

Boam, D., A. Clark, and K. Docherty. 1990. Positive and negative regulation of the insulin gene by multiple trans-acting factors. J Biol. Chem. 265: 8285–8296.

Boshart, M., F. Weber, G. Jahn, K. Dorsch-Hasler, B. Fleckenstein, and W. Schaffner. 1985. A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus. Cell 41: 521–530.

Burglin, T.R., M. Finney, A. Coulson, and G. Ruvkun. 1989. Caenorhabditis elegans has scores of homeobox genes. Nature 341: 239–243.

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.

Cohen, B., M.E. McGuffin, C. Pfeifle, D. Segal, and S.M. Cohen. 1992.apterous, a gene required for imaginal disc development in Drosophila encodes a member of the LIM family of developmental regulatory proteins. Genes & Dev. 6: 715–729.

Corbit, S.R., E. Henderson, H. Masuoka, P.A. Weil, and R. Stein. 1991. Pancreatic β-cell-type-specific transcription of the insulin gene is mediated by basic helix–loop–helix DNA-binding proteins. Mol. Cell. Biol. 11: 1734–1738.

Crowe, D.T. and M.-J. Tsai. 1989. Mutageneis of the rat insulin II 5′-flanking region defines sequences important for expression in HIT cells. Mol. Cell. Biol. 9: 1784–1789.

deWet, J., K.V. Wood, M. DeLuca, D.R. Helsinki, and S. Subramani. 1987. Firefly luciferase gene: structure and expression in mammalian cells. Mol. Cell. Biol. 7: 725–737.

Dong, J., S.L. Asa, and D.I. Drucker. 1991. Islet cell and extra-pancreatic expression of the LIM domain homeobox gene isl-1. Mol. Endocrinol. 5: 1633–1641.

Duprey, P., K. Chowdhury, G.R. Dressler, R. Balling, D. Simon, J-L. Guenet, and P. Gruss. 1988. A mouse gene homologous to the Drosophila gene caudal is expressed in epithelial cells from the embryonic intestine. Genes & Dev. 2: 1647–1654.

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.

Fried, M. and D.M. Crothers. 1981. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. 9: 3047–3060.

Freyd, G., S.K. Kim, and H.H. Horvitz. 1990. Novel cystein-rich motif and homeo domain in the product of the Caenorhabditis elegans cell lineage gene lin-11. Nature 344: 876–879.

Frumkin, A., Z. Rangini, A. Ben-Yehuda, Y. Gruenbaum, and A. Fainsod. 1991. A chicken caudal homologue, CHox-cad, is expressed in the epiblast with posterior localization and in the early endodermal lineage. Development 112: 207–219.

Garner, M.M. and A. Revzin. 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: Application to components of the Escherichia coli lactose operon regulating system. Nucleic Acids Res. 9: 3047–3060.

German, M.S., L.G. Moss, and W.J. Rutter. 1990. Regulation of insulin gene expression by glucose and calcium in transfected primary islet cultures. J. Biol. Chem. 265: 22063–22066.

German, M.S., M.A. Blanar, C. Nelson, L.G. Moss, and W.J. Rutter. 1991. Two related helix-loop-helix proteins participate in separate cell-specific complexes that bind to the insulin enhancer. Mol. Endocrinol. 5: 292–299.

German, M.S., L.G. Moss, J. Wang, and W.J. Rutter. 1992. The insulin and islet amyloid polypeptide genes contain similar cell-specific promoter elements that bind identical β-cell nuclear complexes. Mol. Cell. Biol. 12: 1777–1788.

Giddings, S., F. Chirgwin, and M.A. Permutt. 1985. Evaluation of rat insulin messenger RNA in pancreatic and extra pancreatic tissues. Diabetologia 28: 343–347.

Green, D.W., G. Gomez, and G.H. Greeley. 1989. Gastrointes- tinal peptides. Gastroenterol. Clin. North Am. 18: 695–733.

Hanahan, D. 1985. Heritable formation of pancreatic β-cell tumors in transgenic mice expressing recombinant insulin/simian virus 40 oncogenes. Nature 315: 115–122.

James, R. and J. Kazenwadel. 1991. Homeobox gene expression in the intestinal epithelium of adult mice. J. Biol. Chem. 266: 3246–3251.

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.
German et al.

Acad. Sci. 84: 8819--8823.

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.

Karlsson, O., S. Thor, T. Norberg, H. Ohlsson, and T. Edlund. 1990. Insulin gene enhancer binding protein isl-1 is a member of a novel class of proteins containing both a homeo- and a cys-his domain. Nature 344: 879--882.

Li, P.M., J. Reichert, G. Freyd, H.R. Horvitz, and C.T. Walsh. 1991. The LIM region of a presumptive Caenorhabditis elegans transcription factor is an iron-sulfur- and zinc-containing metalloprotein. Proc. Natl. Acad. Sci. 88: 9210--9213.

Mlodzik, M., and W.J. Gehring. 1987. Expression of the caudal gene in the germ line of Drosophila: Formation of an RNA and protein gradient during early embryogenesis. Cell 48: 465--478.

Nelson, C., V.R. Albert, H.P. Elsholtz, and L.I.-W. Lu. 1988. Activation of cell specific expression of rat growth hormone and prolactin genes by a common transcription factor. Science 239: 1400--1405.

Nelson, C., L.-P. Shen, A. Meister, E. Fodor, and W.J. Rutter. 1990. Pan: A transcriptional regulator that binds chymotrypsin, insulin, and AP-4 enhancer motifs. Genes & Dev. 4: 1035--1043.

Nishi, M., T. Sanke, S. Seino, R.L. Eddy, Y.S. Fan, M.G. Byers, T.B. Shows, G.I. Bell, and D.F. Steiner. 1989. Human islet amyloid polypeptide gene: complete nucleotide sequence, chromosomal location, and evolutionary history. Mol. Endocrinol. 3: 1775--1781.

Ohlsson, H., O. Karlson, and T. Edlund. 1988. A β-cell-specific protein binds to the two major regulatory sequences of the insulin enhancer. Proc. Natl. Acad. Sci. 85: 4228--4231.

Picet, R.L., W.R. Clark, R.H. Williams, and W.J. Rutter. 1972. An ultrastructural analysis of the developing embryonic pancreas. Dev. Biol. 29: 436--467.

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.

Way, J.C. and M. Chalfie. 1988. mec-3, a homeobox-containing gene that specifies differentiation of the touch receptor neurons in C. elegans. Cell 54: 5--16.

Walker, M.D., C.W. Park, A. Rosen, and A. Aronheim. 1990. A cDNA from a mouse pancreatic β-cell encoding a putative transcription factor of the insulin gene. Nucleic Acids Res. 18: 1159--1166.

Walker, M.D., T. Edlund, A.M. Boulet, and W.J. Rutter. 1983. Cell-specific expression controlled by the 5′ flanking regions of the insulin and chymotrypsin genes. Nature 306: 557--581.
Synergistic activation of the insulin gene by a LIM-homeo domain protein and a basic helix-loop-helix protein: building a functional insulin minienhancer complex.

M S German, J Wang, R B Chadwick, et al.

*Genes Dev.* 1992, 6:
Access the most recent version at doi:10.1101/gad.6.11.2165

References
This article cites 48 articles, 23 of which can be accessed free at:
http://genesdev.cshlp.org/content/6/11/2165.full.html#ref-list-1

License

Email Alerting Service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.