

Tissue-specific regulation of the insulin gene by a novel basic helix–loop–helix transcription factor

Francisco J. Naya, Christine M.M. Stellrecht, and Ming-Jer Tsai¹

Department of Cell Biology, Baylor College of Medicine, Houston, Texas 77030 USA

The insulin gene is one of the best paradigms of tissue-specific gene expression. It is developmentally regulated and is expressed exclusively in the pancreatic β -cell. This restricted expression is directed by a tissue-specific enhancer, within the promoter, which contains an E-box sequence. The insulin E-box binds an islet-specific protein complex, termed 3a1. E-boxes bind proteins belonging to the basic helix–loop–helix (bHLH) family of transcription factors. The bHLH proteins function as potent transcriptional activators of tissue-specific genes by forming heterodimers between ubiquitous and cell-restricted family members. In addition, the cell-restricted bHLH members play an important role in specifying cell fate. To isolate the tissue-specific bHLH factor controlling insulin gene expression and study its role in islet cell differentiation, a modified yeast two-hybrid system was utilized to clone a novel bHLH factor, BETA2 (β -cell E-box *trans*-activator 2), from a hamster insulin tumor (HIT) cell cDNA library. Northern analysis demonstrates that high-level expression of the BETA2 gene is restricted to pancreatic α - and β -cell lines. As expected of tissue-specific bHLH members, BETA2 binds to the insulin E-box sequence with high affinity as a heterodimer with the ubiquitous bHLH factor E47. More importantly, antibody supershift experiments clearly show that BETA2 is a component of the native insulin E-box-binding complex. Transient transfection assays demonstrate that the BETA2/E47 heterodimer synergistically interacts with a neighboring β -cell-specific complex to activate an insulin enhancer. In contrast, other bHLH factors such as MyoD and E47, which can bind to the insulin E-box with high affinity, fail to do so. Thus, a unique, cooperative interaction is the basis by which the insulin E-box enhancer discriminates between various bHLH factors to achieve tissue-specific activation of the insulin gene.

[Key Words: Insulin gene; BETA2; bHLH domain; transcription; tissue-specificity; synergism]

Received February 14, 1995; accepted March 7, 1995.

Insulin, a potent growth factor and an essential peptide hormone regulating energy metabolism in vertebrates, is produced exclusively in the β cells of the pancreatic islets of Langerhans. The insulin gene is developmentally regulated by a highly specific mechanism that ultimately restricts the expression of insulin to the pancreatic β cell. The mechanism leading to the cell-specific expression of the insulin gene is complicated and largely unknown. However, it has been firmly established that the cell-specific expression of the insulin gene is directed by 5'-flanking sequences (Walker et al. 1983; Hanahan 1985).

Detailed deletion and mutation analyses have revealed multiple *cis*-acting elements, within the 5'-flanking sequences, which are important for insulin gene transcription (Edlund et al. 1985; Karlsson et al. 1987; Crowe and Tsai 1989; Whelan et al. 1989). One of these elements, RIPE3 (rat insulin promoter element 3), is sufficient to

confer tissue specificity to a non-tissue-specific heterologous promoter in transient transfection assays (Hwung et al. 1990). Further characterization revealed that RIPE3 consists of two separate subelements, RIPE3a and RIPE3b, which function cooperatively to generate maximal activity of the tissue-specific enhancer (Hwung et al. 1990). Notably, each subelement binds β -cell-specific as well as ubiquitous factors (Shieh and Tsai 1991). The RIPE3a site contains the E-box sequence, GCCATC-TGC, which is conserved in all characterized mammalian insulin genes (Steiner et al. 1985). The high degree of conservation strongly suggests the importance of this *cis*-regulatory element. Experimentally, mutation of the E-box has a deleterious effect on insulin enhancer activity (Karlsson et al. 1987; Crowe and Tsai 1989). Thus, it would appear that the E-box-binding proteins are critical for insulin gene expression.

E-boxes bind factors belonging to the basic-helix–loop–helix (bHLH) family of transcription factors (Murre et al. 1989a). The bHLH motif is defined by two amphipathic helices that act as dimerization domains, sepa-

¹Corresponding author.

rated by a nonconserved loop and an adjacent amino-terminal basic region necessary for DNA binding (Murre et al. 1989a; Davis et al. 1990; Voronova and Baltimore 1990). The bHLH proteins are classified into three groups according to their DNA-binding properties, structural features, and tissue distribution. The class A members bind DNA as homodimers or heterodimers and are ubiquitously expressed. The class B members are tissue-specific and appear to homodimerize poorly (Lassar et al. 1991); however, they heterodimerize with the class A members and bind E-boxes with high affinity (Murre et al. 1989b). The class C members are characterized by containing a leucine zipper motif (Landschulz et al. 1988) carboxy-terminal to the bHLH domain and are further categorized into a broadly expressed or tissue-restricted subfamily (Lüscher and Eisenmann 1990). The presence of the leucine zipper appears to determine dimerization specificity (Beckmann and Kadesch 1991) because class C members dimerize exclusively with each other and are unable to heterodimerize with either the class A or class B bHLH factors.

The bHLH family of transcription factors regulate the expression of many cell-specific genes and in most cases, if not all, play a central role in specifying cell fate (Jan and Jan 1993; Weintraub 1993). The best studied system is the process of muscle differentiation induced by myogenic bHLH factors (Olson 1990; Weintraub et al. 1991; Edmondson and Olson 1993; Olson and Klein 1994). Specifically, the activation of the myogenic program and, ultimately, muscle-specific genes, is directed by heterodimers formed between the ubiquitous class A [e.g., E12/47 (Murre et al. 1989a)] and the muscle-specific class B [e.g., MyoD (Davis et al. 1987)] bHLH proteins. In fact, bHLH proteins are associated with diverse developmental programs such as neurogenesis (Villares and Cabrera 1987), hematopoiesis (Mellentin et al. 1989), mesoderm formation (Thisse et al. 1988), and sex determination (Caudy et al. 1988).

There is substantial evidence that members of the bHLH family of transcription factors are components of the insulin E-box-binding complex, termed 3a1. Specific antibodies raised against the ubiquitous bHLH E2A factors, E12/47, recognize proteins in the 3a1 complex (Nelson et al. 1990; Walker et al. 1990; German et al. 1991; Shieh and Tsai 1991). A related class A protein, BETA1 (β -cell E-box *trans*-activator 1), has also been detected in 3a1 (Peyton et al. 1994). BETA1 is the hamster homolog of HeLa E-box-binding factor (HEB) (Hsu et al. 1992). Previous studies have shown that the 3a1 complex is present solely in pancreatic α - and β -cell line nuclear extracts (Ohlsson et al. 1988, 1991; Aronheim et al. 1991; Shieh and Tsai 1991; Robinson et al. 1994). Because class A bHLH proteins are expressed ubiquitously and almost exclusively form heterodimers with class B factors, it is unlikely that they alone mediate the tissue specificity of the insulin E-box complex. A reasonable scenario is that pancreatic β cells contain an islet-specific bHLH factor, as has been proposed by several investigators (Cordle et al. 1991; German et al. 1991; Shieh and Tsai 1991; Park and Walker 1992). This putative

tissue-specific factor presumably associates with one or more of the ubiquitous bHLH proteins, thereby mediating the restricted expression of the insulin gene.

Despite intensive efforts by several laboratories to clone the putative islet-specific bHLH factor using conventional expression screening procedures, only class A factors have been isolated. This is not unexpected because the ubiquitous bHLH proteins form homodimers and bind to E-boxes *in vitro* with high affinity. In contrast, the tissue-specific bHLH factors fail to bind to DNA as homodimers. Thus, the expression screening approach has been unsuccessful in cloning the islet-specific bHLH factor. For this reason, we have taken advantage of the functional properties of bHLH factors and used a modified yeast two-hybrid system (Fields and Song 1989) in an attempt to clone the putative islet-specific factor. In this paper we report the isolation of BETA2, a novel pancreatic α - and β -cell specific bHLH transcription factor. BETA2 binds the insulin E-box sequence with high affinity as a heterodimer with E47 in a gel shift assay. Antibody supershift experiments demonstrate that BETA2 is a component of the native insulin E-box complex. In addition, the BETA2/E47 heterodimer synergistically interacts with an adjacent β -cell-specific complex, 3b, to activate the insulin gene in a tissue-specific manner.

Results

Screening strategy and isolation of BETA2

To clone cDNAs that encode proteins that interact with the ubiquitous bHLH factor, E47, and bind the insulin E-box we used a modified yeast two-hybrid system (Fields and Song 1989). The reporter used for this two-hybrid screen contained a single copy of the insulin E-box sequence inserted upstream of a yeast *iso-1-cytochrome c* (*cyc1-P*) minimal promoter and the *Escherichia coli lacZ* gene. Because we were looking for bHLH proteins specifically expressed in pancreatic β -cells, a hamster insulin tumor (HIT) cell cDNA library was constructed as a fusion with the GAL4 activation domain. The third component of this system was a vector designed to express only the DNA-binding and dimerization domain (i.e., bHLH region) of E47, thus lacking transcriptional activity (Quong et al. 1993). Therefore, those fusion proteins that can dimerize with the E47 bHLH domain and bind to the insulin E-box sequence will activate the reporter and turn the host yeast cells blue on X-gal plates (Fig. 1a).

An oligo(dT)-primed HIT cell cDNA plasmid library consisting of 500,000 bacterial recombinants was transformed into the protease-deficient yeast strain BJ2168, carrying the reporter and E47 bHLH expression plasmids and plated out on selective media containing X-gal. A total of 500,000 yeast colonies were screened and only one clone, 11B-1, demonstrated properties expected of a tissue-specific bHLH member and was renamed BETA2. As shown in Figure 1b, BETA2, together with E47, acti-

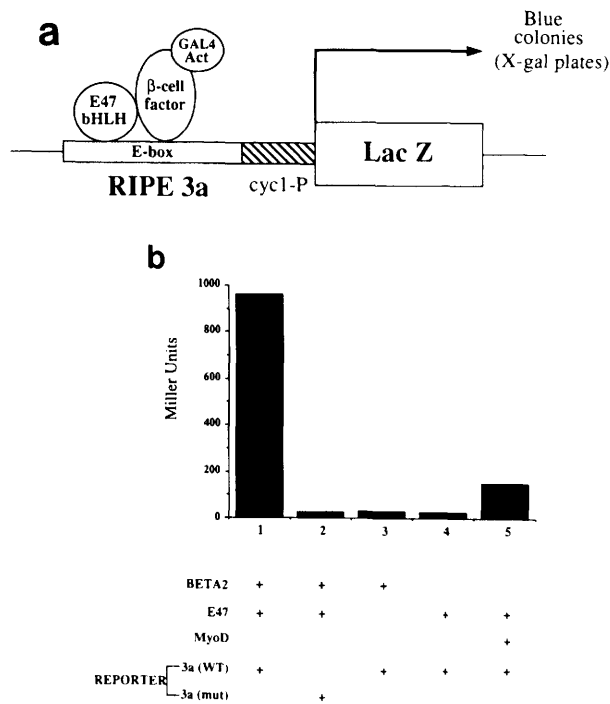


Figure 1. Screening strategy for isolation of bHLH cDNA clones and β -galactosidase assay results. (a) A schematic representation of a modified yeast two-hybrid system. The reporter contains the insulin E-box sequence, termed RIPE3a, inserted upstream of the yeast iso-1-cytochrome *c* minimal promoter (*cyc1-P*) and the *E. coli lacZ* gene. Interaction between bHLH-GAL4 fusion products and E47 bHLH region with subsequent binding to the E-box will activate the reporter and yield blue colonies when plated on media containing X-gal. (b) β -Galactosidase activity of the BETA2-GAL4 activation domain fusion in the protease-deficient yeast strain BJ2168. Also depicted are the activities of various bHLH expression plasmids in yeast using the RIPE3a reporter. Activities are expressed as Miller units (OD/mg of protein per min). Values given are the average of at least three independent transformations.

vates the E-box reporter (lane 1) but fails to activate a mutant reporter (lane 2) containing an E-box point mutation that drastically decreases binding and *trans*-activation function of the element. Very little activity was observed with BETA2 (lane 3) or E47 (lane 4) alone. The positive control was a MyoD-GAL4 activation domain fusion (Staudinger et al. 1993), E47, and reporter plasmids (lane 5). Previous studies from our laboratory have shown that MyoD, a muscle-specific bHLH factor, heterodimerizes with E47 and binds to the insulin E-box sequence with high affinity (M.J. Peyton and M.-J. Tsai, unpubl.).

DNA sequence analysis of the 2.5-kb BETA2 cDNA revealed an open reading frame of 1143 nucleotides. The first methionine (Fig. 2a, bold and circled) is in-frame with the GAL4 activation domain in the cDNA library expression vector and is not preceded by an in-frame stop codon. The open reading frame encodes a protein of 381

amino acids with a predicted size of 42 kD. The deduced amino acid sequence of BETA2 reveals an amino-terminal bHLH domain between residues 126 and 181 (Fig. 2a, underlined). Alignment of the BETA2 bHLH domain with other members of the bHLH family indicates that it belongs to the tissue-specific class B bHLH subgroup (Fig. 2b) and therefore represents a likely candidate as a key regulator of islet cell differentiation. Significantly, BETA2 contains a tyrosine residue within helix II that is absolutely conserved among class B bHLH members, whereas class A members contain a valine residue in the same position, further indicating that this factor is a bona fide member of the tissue-specific subclass. Within the bHLH domain, BETA2 shares 55% amino acid identity with another novel bHLH member isolated from HIT cells, designated BETA3 (M.J. Peyton, C.M.M. Stellrecht, F.J. Naya, H.-P. Huang, and M.J. Tsai, in prep.) and shares 54% amino acid identity with the *Drosophila* neural precursor gene, *atonal* (Fig. 2b, Jarman et al. 1993). Immediately amino-terminal to the bHLH domain of BETA2 is a 54-amino-acid cluster rich in glutamate and aspartate residues (Fig. 2a, bold) followed by several basic amino acids (Fig. 2a, dashed underline). The highly acidic nature of this cluster of amino acids is characteristic of transcriptional activators (Ptashne 1988).

BETA2 mRNA expression

Northern analysis was performed to determine the size, cell type, and tissue distribution of BETA2 transcripts. Northern blots containing total RNA isolated from several endocrine and exocrine pancreas-derived cell lines, pituitary cell lines, and adult hamster tissues were probed with the BETA2 cDNA. The results presented in Figure 3 demonstrate that the BETA2 cDNA hybridizes with a single, prominent 2.6-kb mRNA transcript and a less abundant 1.6-kb transcript. High-level expression of BETA2 mRNA transcripts was observed in a pancreatic α -cell line (α TC) and pancreatic β -cell lines [(HIT) hamster; (β TC) mouse; and (RIN) rat] but not in the pancreatic acinar cell line (AR42J, rat) or in the pancreatic ductal cell lines (MDAPanc-3, human; BxPC3, human), demonstrating that this bHLH factor is specific to the endocrine cells of the pancreas. Abundant levels of the 2.6-kb transcript are also detected in the brain. A weak hybridizing signal appears in the mouse intestinal enteroendocrine S-cell line (STC-1). Upon longer exposure, a weak signal is also detected in the mouse corticotroph pituitary cell line (AtT20) and hamster intestine (data not shown). No signal is evident in total RNA from hamster salivary gland, testes, lung, kidney, spleen, liver, and whole pancreas. The lack of a signal from the whole pancreas is likely attributable to a dilution effect of the islet cell RNA by an excess of the acinar cell RNA and to the partial degradation of the pancreas RNA samples. In addition, the hamster fibroblast cell line (CHO) and the rat somatotroph pituitary cell line (GH3) do not contain a transcript for BETA2. An 18S rRNA cDNA probe was used as a control for loading and transfer efficiency.

Naya et al.

a

```

1  AGAAGCCAAAGTGTCCCGAGGCTCCGGGGTTATGAGATCGTCACTGCTCAGGACATCTTA  60
  R R Q S V P R L R L R G Y E I V T A Q D I L
61  ACAACAGGAAGTGGAAACATGCAACAAATCATACAGCGAGGCGGAGCTGATGGTGAGCCT  120
  T T G S G N (M) T K S Y S E S G L M G E F
121  CAGCCCAAGGTCCCCAGCTGGACAGCAGAGTGTCTTAGTTCTCAGGACGAGGATCAT  180
  P Q Q G P P S W T D E C L S S Q D E D H
181  GAGCGACAGAAGAGGAGGAGGCTCGAAGCCATGAATGCAGAGGAGGACTCTCTGAGA  240
  E A D K K E D E L E A M N A E E D S L R
241  AACGGAGGAGTGGAGGAGTGAAGTGAAGACCTGGAAGAGGAAGACGAGGAAGAGGAG  300
  N G G D E E D E D E D L E E E D E E E
301  GAAGATGATCAAAAGCCCAAGAGCGGGCCCAAAAAGAAAAGATGACCAAGCGCCG  360
  E D D Q K P K R R G P F K K K M T K A R
361  CTAGAGCGTTTAAATTAAGGCGCATGAGGCCAACCGCCGGGAACCGCAACCGCATG  420
  L E R F K L R R M K A N A R E R N R M H
421  GGGCTGAATGCGAGCTGGACAACCTGCGCAAGTGGTACCTGCTACTCCAAGACCCAG  480
  G L N A A L D N L R K V V V P C Y S K T Q
481  AAGCTCCCAAGATGAGACACTGGCTGGCCCAAAATACATTTGGGCTCTGTGAGAG  540
  K L S K I E T L R L A K N Y I W A L S E
541  ATCTCGGCTCAGGCAAAAGCCCTGACCTGGTCTCTCTCGTTCAAACGCTCTGCAAAAG  600
  I L R R S G K S P D L V S F V Q T L C K G
601  TTGCGCCAGCCACCACCACTGGTGTGCTGGCTGCTGCAACTCAATCTCGGACTTTTC  660
  L S Q P T T N L V A G C L Q L N P R T F
661  CTGCGCAAGACAGAACCCGAGATGCCCCCGCATCTGCCACAGCCAGTGTCTTCCCG  720
  L P E Q N P D M P P H L P T A S A S F P
721  GTACACCCCTACTCCTACAGTCCCCGGGCTACCGAGCCCTCCTATGGCACCATGGAC  780
  V H P Y S Y Q S P G L P S P P Y G T M D
781  AGCTCCCATGTCTTCCAAAGTGAAGCGCCGCCACATGCTACAGCGGACTTTGGAACCC  840
  S S H V F Q V K P P P H A Y S A T L E P
841  TTCTTTGAAAGCCCTGACTGATGACACGACCCCTTTTGTATGGACCCCTCAGCCCG  900
  F F E S P L T D C T S P S F D G P L S P
901  CCGCTCAGCATCAATGGCACTTCTCTTTCAAACACGAAACATCCCGCGGACTTTGAAA  960
  P L S I N G N F S F K H E P S A E F E K
961  AATTATGCCCTTACCATGCACTACCTGCGAGCAGCCCTGGCAGGGCCCAAGCCAGGA  1020
  N Y A F T M H Y P A A T L A G P Q S H G
1021  TCAATCTCTCGGGTGCACCTGCCCCCTGCTGCGAGATCCCCATAGACAATATATGCT  1080
  S I F S G A T A P R C E I P I D N I M S
1081  TTGATAGCCATTCACATCATGAGCGAGTCTGATGAGTCCGCGGACTCAATGCCATTTT  1140
  F D S H S H E R V M S A Q L N A I F H
1141  GATTAGGACCGTCACTTCCACATTCGCGGAAACGAACCACTGTGCTTTCAGTGAC  1200
  D *
1201  TGCTCGTGTTCACAGAAGCGCCCTTTTGTACGATTTCTGCAAGTGAATACTCAAA  1260
  GCTTCAAGTGATATGATTTATTTGTCGTACTGCTTTGGAAGAAACAGGGGATCAAA
1261  GCTCCTGTTCACCTTATGATTTGTTCTATAGCTCTCTATTTTAAAAAATAAACAGT  1380
  AAAGTAAAAAAGAAATGTTGACCCAGAAATTTGCTGTAGCTGTACTCAGATCATATTA
1381  TATCTGATCGGGATAAAAAATCACAAGCAATAATAGGATCTATGCAATTTTAAACTA  1440
  GTAATGGCCCAATTAATAATATATAAATAATATATTTTCAACCAGCAATTTACTACTG
1441  TTACCTTTCCATGCTGAATTTTGTGTGATTTTGTACAGAAATTTTAAAGACATTTT  1500
  TATAACGTGGATTTCTTATTTTAAACCAATGACGCTCATCAATTTTATACATATCAGA
1461  AAAGTGAATATATCTAATTTATACAAAGATAATTAACATAATTTAAACAGCAGAAAG  1680
  TGCTTAGAAGTTATTCGCTTGCCTTACGACTCTCTCTCTCAATTTAAACAGCAGAAAG
1741  AACAAAACAAAACAAATTCACAAATTTGAGCAATCCATTTCACTTTGAAGTTGTCCTC  1860
  TCCTTAAAAATAGAAACAGATTCATAGTATTCAGAGGATGAAAAAGAGATGCTTCTCT
1861  ATCAAAACATATCAATTCATCTACTTGCACAGCTTGTATATACATATTAATAACAAATG  1920
  CAACATACACACACACTCTCTTAAATCAAAGCTGTGACTATACACACAAATTTGCAC
1981  CTCTTCAATTTTGTCTTTCTTTGAATTCAGATTTTATGAGTGTCTGCACACTGTG  2040
  ATGTTCCAGAAAATATAAATGCAATTTTATACATAGTCAAAACAAATGTTGGCTGTG  2100
  ATCTATTATGTAATAAATCTGAGCCTAAAATTTTACAGGTGTTAAAGTTGGGATGAT  2160
  ACCTAGCTAGTCACTAGTACAGTACTTAAATTAATCAAACCTTGAATCCATAAT  2280
  AGAACACATAGCTATTGATGTAAGAAAGTCAAGTCCAGAACAGTGTGCTCAGGATGGAT  2340
  GCTGTCACCTGAAATGACCTGTACTGTAATCTGTTTGTACCCCTGTGATATGTTG  2400
  GTAACGCACAAATTTGAAAACGCTCATGCGATGCCAATAAAAATAGTATTGAAAATCAA  2460
  AAAAAAAAAAAAAAAAAA 2476

```

b

```

          BASIC          HELIX I          HELIX II
consensus  --R--N-RER-R--Ψ  F--L-----KW-IL--AW-YP  L
          T          L          T          V          V
BETA2      LRRMKANARERNRMHGLNALDNLKRVVP.CYSKTKQLSKIETRLRLAIYIWLSEI
atonal     K--LA-----R  QN--Q-F--R--Q-L-  -LGNDRQ--H--QM-QT--S--GDL
tal-2      T-KIFT-T--W-QQNV-S-FAK--LL-  THPPDK----N-----MR--NF-KVK
tal-1(scl) V--IFT-S---W-QQNV-F-FAE--DLI-.  THPPDK----N-I----MK--NF-AKL
lyl        A--VFTNS---W-QQNV-G-FAE--LL-.  THPPKR----N-VL---MK--GF-VRL
MASH-2    A--...E-----VKLV-LGFQA--QH--.  HGGANK-L--V---S-VE--R--QRL
MASH-1    A--...E-----VKLV-LGFAT--EH--.  NGAANK-M--V---S-VE--R--QQL
xtwtist   SQ-VM--V--Q-TQS--E-FSS--II-.  TLPSD- ----Q--K--SR--DF-CQV
mtwtist   TQ-VM--V--Q-TQS--E-RAA--II-.  TLPSD- ----Q--K--AR--DF-YQV
MyoD      D--KA-TM--R-R-LSKV-E-FET-KRCTS  SNF-N-R-P-V-I--N-IR--EG-QAL

```

Figure 2. BETA2 cDNA sequence and alignment of bHLH domain of class B factors. (a) The deduced amino acid sequence is shown as single-letter code. The first in-frame methionine is indicated as bold and circled. The open reading frame continues through nucleotide 1143, and the stop codon is indicated by an asterisk. The bHLH domain is underlined. This is preceded by acidic (bold) and basic (dashed underline) residues. (b) Alignment of BETA2 bHLH domain with other tissue-specific bHLH proteins. Consensus bHLH residues are from Murre et al. (1989a). The bHLH domain of BETA2 is aligned with the bHLH domain of class B factors from various species: atonal (Jarman et al. 1993); tal-2 (Xia et al. 1991); tal-1 (scl) (Begley et al. 1989); lyl (Mellentin et al. 1989); MASH-2 and MASH-1 (Johnson et al. 1990); xtwtist (Hopwood et al. 1989); mtwtist (Wolf et al. 1991); MyoD (Davis et al. 1987). Dashes denote sequence identity. (Ψ) Hydrophobic residues.

DNA-binding properties of BETA2

To verify that BETA2 can bind to the insulin E-box sequence we examined the DNA-binding properties using a gel mobility shift assay. Because the mRNA size indicated that we had isolated a nearly full-length cDNA, the entire 2.5-kb cDNA of BETA2 was transcribed and translated *in vitro*. To examine the fidelity of the translated products reactions were carried out in the presence of [³⁵S] methionine. Equimolar amounts, as determined by SDS-PAGE, of separately translated unlabeled BETA2 and shPan1 (full-length hamster homolog of E47) were mixed and incubated in the presence of a radiolabeled insulin E-box oligonucleotide and analyzed by gel mobility shift assay. This reaction generated two complexes, a slower migrating shPan1 homodimer (designated as Pan1) and a faster migrating complex suggesting the formation of a heterodimer as indicated (Fig. 4, lane 4). An adjacent lane shows the Pan1 homodimer (lane 2). As

expected of tissue-specific bHLH members, *in vitro*-translated BETA2 does not bind the insulin E-box sequence alone (lane 3). These results are consistent with the property that the tissue-specific bHLH proteins predominantly bind E-box sequences with high affinity as heterodimers with a ubiquitous bHLH partner (Murre et al. 1989b). The binding to the E-box sequence is specific in that an excess of unlabeled oligonucleotide (mut 1), containing an E-box point mutation that does not affect binding of the insulin E-box complex or promoter activity (Whelan et al. 1990), effectively competes for the BETA2/Pan1 heterodimer (lanes 5, 6). In contrast, an oligonucleotide (mut 2) containing a point mutation that abolishes insulin E-box binding and promoter activity (Whelan et al. 1990) is unable to compete for the BETA2/Pan1 heterodimer (lanes 7, 8) revealing similar binding characteristics with the native factor.

To determine whether BETA2 participates in the tissue-specific insulin E-box-binding complex, termed 3a1,

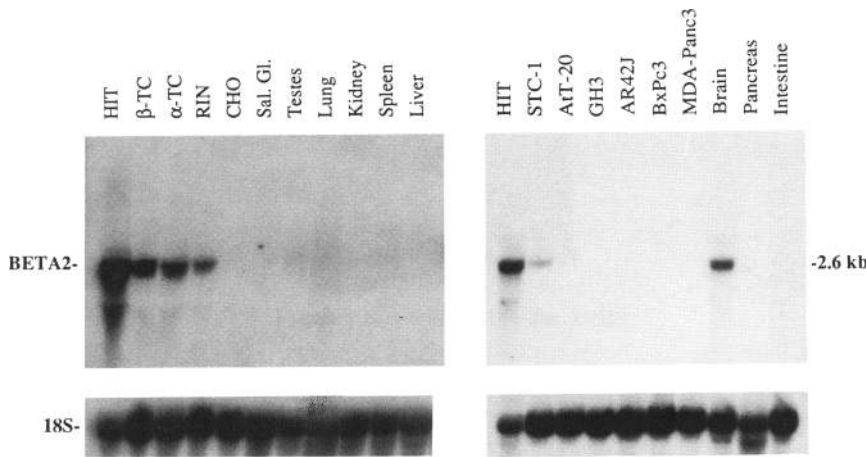


Figure 3. Northern blot analysis of tissue and cell distribution of BETA2 transcripts. Lanes contain 10 μ g of total RNA isolated from various cell lines and selected hamster tissues. The blot was probed with random-primed BETA2 cDNA. An \sim 2.6-kb transcript is observed in HIT, β -TC, α -TC, RIN, STC-1, AtT20, brain, and intestine. The RNA blot was reprobed with an 18S rRNA cDNA to check loading and transfer efficiency.

antibodies were raised against the carboxy-terminal region of BETA2. The specificity of these antibodies was first tested against in vitro-translated BETA2/Pan1 heterodimers using a gel shift assay. The BETA2 antibodies specifically recognize the heterodimer and supershift the complex to a slower migrating position (Fig. 4, lanes 9, 10; asterisks). In contrast, the BETA2-specific antibodies do not recognize the Pan1 homodimer (cf. lanes 9 and 10 with lanes 11 and 12). Also, the preimmune sera has no effect on either the homodimer or the heterodimer (lanes 11, 12). The BETA2 antibodies were subsequently tested against HIT cell nuclear extract. The presence of the 3a1 complex was verified by specific competition using the mut 1 (Fig. 5, lanes 8, 9) and not with the mut 2 E-box oligonucleotides (lanes 10, 11). Incubation of the HIT nuclear extract with the BETA2 antibodies results in the disappearance of the 3a1 band (lanes 2–4) and the appearance of a slower migrating supershifted complex (lanes 2–4; asterisks). An excess of preimmune sera has no effect on the 3a1 complex (lanes 6, 7). These results confirm that BETA2 is a component of the insulin E-box complex, 3a1.

Transcriptional activation by BETA2

To examine the activation potential of BETA2 on the insulin E-box enhancer, transient transfection assays were carried out in the insulin-producing cell line HIT. HIT cells were cotransfected with mammalian expression vectors containing the entire BETA2 cDNA and the full-length E47 cDNA (Henthorn et al. 1990). The insulin E-box constructs RIPE3(WT) and RIPE3(M) were used as the reporters. The wild-type reporter RIPE3(WT) contains three copies of the insulin promoter nucleotide sequence -126 to -86 inserted at the -50 site of the ovalbumin minimal promoter chloramphenicol acetyltransferase (CAT) construct pOVCAT-50 (Hwung et al. 1990). This enhancer element contains the insulin E-box sequence 3a and an adjacent β -cell-specific-binding site, 3b. The mutant reporter RIPE3(M) contains an E-box

point mutation that drastically decreases the *trans*-activation function of the enhancer (Shieh and Tsai 1991).

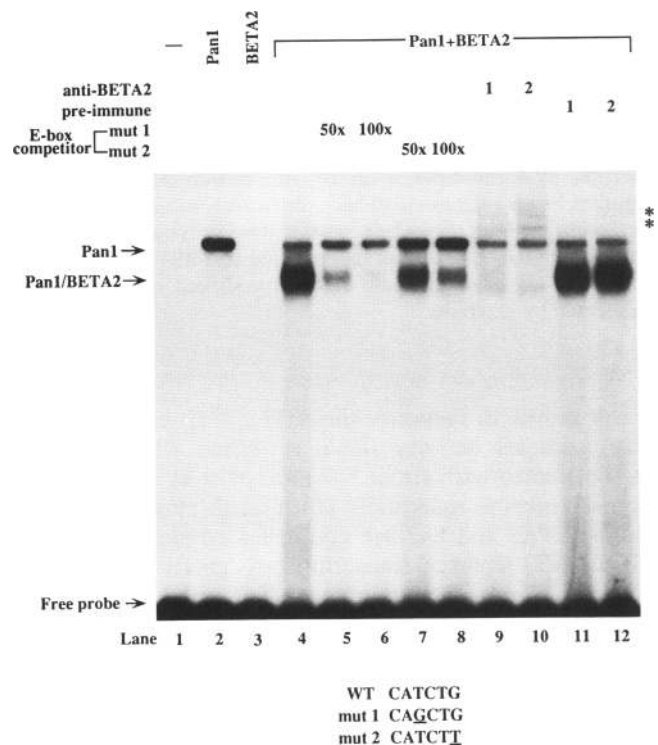


Figure 4. DNA-binding properties of BETA2 and anti-BETA2 antibody specificity. In vitro-translated products were assayed by gel mobility shift for ability to bind to the insulin E-box containing RIPE3 probe. Binding specificity was determined by competition with 50- and 100-fold molar excess of unlabeled double-stranded mutant E-box oligonucleotides. E-box sequences of mutant oligonucleotides are shown with mutant nucleotides underlined. The Pan1 and Pan1/BETA2 complexes are indicated. One microliter of each translated product was used for the binding reactions. Antibody specificity was assessed by adding 1 or 2 μ l of IgG-purified anti-BETA2 antibody or preimmune sera after completion of binding reactions. Supershifted complexes are indicated by asterisks.

Naya et al.

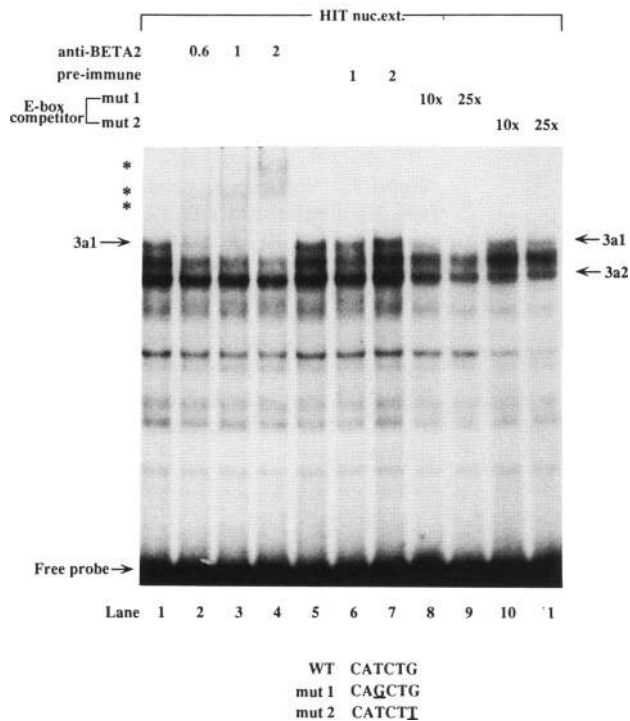


Figure 5. BETA2 participates in the insulin gene E-box-binding complex. HIT nuclear extract was assayed by gel mobility shift. The location of the tissue-specific E-box-binding complex 3a1 is verified by competition with 10- and 25-fold molar excess of mutant E-box oligonucleotides as indicated. Binding reactions, containing 3.5 μ g of nuclear extract, were initiated prior to the addition of 0.6, 1, or 2 μ l of IgG-purified anti-BETA2 antibodies or 1 to 2 μ l of preimmune sera. Supershifted complexes are indicated by asterisks.

As shown in Figure 6, the RIPE3(WT) reporter alone has minimal activity (lane 1). When RIPE3(WT) is cotransfected with either E47 (lane 3) or BETA2 (lane 5), transcription is marginally stimulated. However, when both BETA2 and E47 are cotransfected along with the RIPE3(WT) reporter, transcription is greatly stimulated (lane 7). This activity is dependent on the E-box sequence because the BETA2/E47 heterodimer does not activate the mutant RIPE3(M) reporter (lane 8). These results clearly demonstrate that the BETA2/E47 heterodimer functions as a potent *trans*-activator of the RIPE3 enhancer in insulin-producing cells.

For the next series of experiments we addressed the paradox why MyoD/E47 heterodimers or E47/E47 homodimers, which bind the insulin E-box with high affinity *in vitro*, are not functionally competent *in vivo* to activate the insulin gene either in β cells or in other cell types. As described previously, RIPE3 contains both the 3a and 3b tissue-specific binding sites. It has been established that 3a and 3b synergize to reach the maximum enhancer activity of RIPE3 (Hwung et al. 1990). Therefore, it is possible that the RIPE3 enhancer may utilize synergism to discriminate among various bHLH factors. If this is the case, we would expect that only BETA2 is

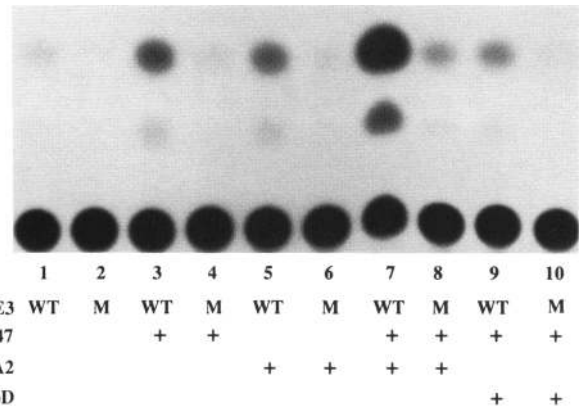


Figure 6. BETA2 activates the RIPE3 enhancer in HIT cells. HIT cells transfected as in Fig. 7 except the wild-type and mutant RIPE3 reporters were used. RIPE3 (M) contains an E-box point mutation that has been demonstrated previously to drastically reduce the *trans*-activation potential of the element (Shieh and Tsai 1991). These are representative data of at least three independent transfections. CAT activity was determined for the above samples and incorporated with the values given in Fig. 7.

functionally competent to interact with the neighboring 3b factors. To investigate this question, transient transfection assays were carried out in two different mammalian cell lines using various insulin E-box reporters. The noninsulin-producing fibroblast cell line BHK and the insulin-producing cell line HIT were cotransfected with mammalian expression vectors containing the entire BETA2 cDNA, the full-length E47 cDNA, or the entire MyoD cDNA (Davis et al. 1987). The insulin E-box construct, RIPE3, was used as a reporter. As a control we used the RIPE3a reporter, which contains four copies of the insulin promoter nucleotide sequence -110 to -86 (Hwung et al. 1990). This construct contains the 3a- but not the 3b-binding site.

The results presented in Figure 7 summarize the transient transfection experiments. Using the RIPE3a re-

Cell line	BETA2/E47	MyoD/E47	E47/E47
(3a) HIT	0.03 \pm 0.08	0 \pm 0.02	0.13 \pm 0.12
(3b) (3a) BHK	0.95 \pm 0.55	1.74 \pm 0.42	0.37 \pm 0.19
(3b) (3a) HIT	20.32 \pm 9.54	1.67 \pm 0.03	1.72 \pm 0.84

Figure 7. Activation of insulin enhancer constructs by various bHLH factors in HIT and BHK cells. Noninsulin-producing BHK and insulin-producing HIT cells were transfected with the RIPE3 or RIPE3a reporters, together with either of the various expression plasmids, E47, BETA2, and MyoD. CAT enzyme activity was measured in cell extracts 48 hr. later. Values shown are averages of percent conversion of [14 C]chloramphenicol after subtracting basal reporter activity of at least three independent transfections \pm S.D.

porter, the BETA2/E47 heterodimer, the MyoD/E47 heterodimer, and the E47/E47 homodimer exhibit minimal but nearly equal levels of activity in HIT cells. Although the 3b factors are present in these cells, they cannot participate in the BETA2-mediated activation of this reporter lacking the 3b-binding site. Not surprisingly, similar results are obtained when using the entire RIPE3 enhancer in noninsulin-producing BHK cells. In this series of experiments the 3b-binding site is present; however, BHK cells do not contain the 3b-binding complex (Shieh and Tsai 1991). In sharp contrast, a striking difference in transcriptional activity is observed between the various bHLH complexes when the RIPE3 enhancer is used in the insulin-producing HIT cell line. The RIPE3 enhancer is greatly activated by the BETA2/E47 heterodimer as predicted (Fig. 6, lane 7; Fig. 7). Here, the BETA2/E47 heterodimer can functionally and specifically interact with the neighboring 3b-binding complex to activate the insulin enhancer. However, the MyoD/E47 heterodimer (Fig. 6, lane 9; Fig. 7) or the E47/E47 homodimer (Fig. 6, lane 3; Fig. 7) fail to synergize with 3b to activate the reporter. These results clearly demonstrate that the insulin E-box enhancer functionally discriminates between various bHLH factors by utilizing synergism between two distinct classes of tissue-specific DNA-binding proteins. Furthermore, these results confirm that BETA2 is the authentic component of the insulin E-box complex, 3a1, which can cooperate with 3b to exert maximum insulin enhancer activity.

Discussion

We have successfully exploited a modified yeast two-hybrid system to isolate a novel member of the bHLH family of transcription factors important for the regulation of the insulin gene. The pancreatic islet cell specific bHLH factor, BETA2, exhibits properties characteristic of a tissue-specific bHLH family member. BETA2 heterodimerizes with the ubiquitous bHLH factor, E47, and binds to the insulin E-box sequence with high affinity. More important, the antibody supershift experiments show that BETA2 is the authentic pancreatic islet-cell-specific bHLH activator in the insulin E-box complex. Finally, the BETA2/E47 heterodimer synergistically interacts with another β -cell-specific factor to *trans*-activate the insulin E-box enhancer in a tissue-specific manner.

For our purposes, this modified interaction cloning approach was expected to be more specific than the conventional two-hybrid system because sequence-specific DNA binding is required in addition to protein-protein interactions. First, we were looking for cDNAs that interacted with the ubiquitous bHLH factor, E47. Second, these dimerized proteins were then required to bind to the insulin E-box sequence (i.e., the RIPE3a site). This approach is essentially similar to the screening method used to isolate the serum response factor (SRF) accessory protein SAP-1 (Dalton and Treisman 1992). The two major advantages of using our approach over the conventional two-hybrid system is that we could avoid isolating

any negative bHLH regulators that inhibit DNA binding, such as Id (Benezra et al. 1990), and any other proteins that interact with the bHLH domain of E47 specifically but do not bind to the insulin E-box sequence.

Our data suggest that BETA2 is nearly a full-length cDNA clone, as the 2476-nucleotide BETA2 cDNA is sufficiently close in size to the estimated 2.6-kb message observed by Northern analysis. BETA2 contains a 381-amino-acid open reading frame with a calculated size of 42 kD, yet it migrates with an estimated size of 60 kD when the full-length cDNA is translated *in vitro* and resolved by SDS-PAGE. Surprisingly, we also observe a 60-kD protein product when we translate *in vitro* a truncated BETA2 cDNA containing only the open reading frame (nucleotides 1–1143). The overall charge distribution and amino acid content of BETA2 may account for the slower mobility of the protein under SDS-PAGE conditions. As expected, a similar size band is also detected by Western analysis of HIT cell nuclear extracts (data not shown). Interestingly, it has been reported that a putative β -cell-specific bHLH factor was estimated to be a 25-kD protein when the insulin E-box complex is analyzed by an *in vitro* subunit exchange assay (Park and Walker 1992). Although this method is not very accurate, we cannot exclude the possibility that several related islet cell-specific bHLH family members, analogous to the myogenic subfamily of bHLH transcription factors, may exist in pancreatic β cells. We have isolated another novel tissue-restricted bHLH member expressed in HIT cells (M.J. Peyton, C.M.M. Stellrecht, F.J. Naya, H.-P. Huang, and M.-J. Tsai, *in prep.*). The role of this factor in insulin gene regulation is not known at present, as it cannot bind to the insulin E-box by itself or with E47.

The bHLH domain of BETA2 is located in the amino-terminal region of the protein adjacent to a cluster of acidic amino acids. The location of the bHLH domain is unexpected because the majority of the class B factors have the bHLH domain located predominantly towards the carboxy-terminal region. Because the bHLH domain is only ~60 amino acids the remainder of the protein may contain an, uncharacterized, as of yet, tissue-specific and/or functional domain.

Our Northern analysis results demonstrate that the BETA2 gene is expressed abundantly in pancreatic α - and β -cell lines. Coincidentally, it has been shown that the insulin E-box-binding activity 3a1 is present in pancreatic α - and β -cell line nuclear extracts (Ohlsson et al. 1991; Robinson et al. 1994). The expression of BETA2 transcripts in α and β cells, which are believed to originate from a common progenitor (Alpert et al. 1988; Gittes and Rutter 1992; Teitelman et al. 1993), would suggest that BETA2 likely plays a role in the specification of islet cell types in addition to its function as a cell-specific transcriptional activator of the insulin gene. This is consistent with the notion that bHLH factors have an important function in both *trans*-activation and cell-type determination (Weintraub et al. 1991; Jan and Jan 1993; Weintraub 1993; Olson and Klein 1994). BETA2 gene expression is also observed in the brain, intestine,

and in an intestinal endocrine cell line (STC-1). These results are striking from the perspective of islet ontogeny. The presence of BETA2 mRNA transcripts in the intestine agrees with the generally accepted observation that the pancreatic islets are derived from the gut endoderm (LeDouarin 1988; Guz et al. 1995). However, there is evidence that the islets are derived from the ectoderm via neural crest cells (Pearse and Polak 1971). Tyrosine hydroxylase, a neuronal-specific marker, is expressed in the islets (Teitelman et al. 1981; Teitelman and Lee 1987) while transgenes under the control of the rat insulin II promoter are expressed in the developing nervous system (Alpert et al. 1988). Furthermore, insulin transcripts have been detected in mouse brains starting at day 9.5 embryos (Deltour et al. 1993). On the basis of these observations we can speculate that BETA2 acts as a *trans*-activator of the insulin gene in the developing brain; however, its function in the adult brain is unclear. The presence of BETA2 transcripts in tissues of endodermal and ectodermal origin is intriguing and would not immediately appear to shed light on the unresolved issue of islet ontogeny.

The transient transfection assays in the insulin-producing HIT cells and the noninsulin-producing fibroblast BHK cells dramatically illustrate the complexity of the RIPE3 enhancer. First, other tissue-specific bHLH factors, such as MyoD, cannot merely replace BETA2 because only the BETA2/E47 heterodimer functions as a potent *trans*-activator of the RIPE3 reporter in HIT cells. Second, the limited capacity of the BETA2/E47 heterodimer to *trans*-activate the RIPE3 reporter in BHK cells is likely attributable to the absence of important β -cell-specific activators such as 3b. These conclusions are further supported by the observation that the BETA2/E47 heterodimer does not significantly stimulate the transcriptional activity of the RIPE3a reporter above basal levels in HIT cells. The RIPE3a reporter does not contain the 3b site, and, as a result, the factors that bind to this element are unable to participate in the activation of this reporter.

Taken together, our results demonstrate that the tissue-specific activation of the insulin gene requires high-affinity DNA binding by a pancreatic islet cell-specific bHLH factor, BETA2, and synergistic interaction with a neighboring β -cell-specific complex 3b. The reason why the MyoD/E47 heterodimer or the E47/E47 homodimer, which can bind to the insulin E-box with high affinity *in vitro*, fails to activate the RIPE3 enhancer is because of its inability to synergize with 3b. Until recently, the mechanism by which E-boxes functionally discriminate between various bHLH factors *in vivo* has not been characterized sufficiently. A report by Weintraub et al. (1994) describes the identification of a negative control *cis*-acting element in the immunoglobulin heavy chain (IgH) enhancer that prevents MyoD from functioning as an activator of this gene. Similarly, the muscle-specific muscle creatine kinase (MCK) enhancer contains a negative *cis*-acting element that inhibits the activity of the ubiquitous bHLH factor E12. This negative control is directed at the bHLH domains of MyoD and E12, respec-

tively. Our findings represent a unique, cooperative mechanism utilized by the insulin E-box enhancer to discriminate between various bHLH factors to activate the insulin gene. These results would indicate that BETA2 is the authentic tissue-specific bHLH regulator of the insulin gene and that it contains functional domains, as yet unidentified, necessary for the synergistic interaction between neighboring β -cell-specific factors.

Materials and methods

Yeast plasmids and strains

The yeast reporter plasmid containing the *URA3* gene as a selection marker, YRpRIPE3a, was constructed by insertion of an annealed complementary oligonucleotide containing a single copy of RIPE3a into the *Bgl*III site of the vector YRpPC3 (Pham et al. 1991). The *Bgl*III site is directly upstream of the yeast iso-1-cytochrome *c* minimal promoter fused to the *E. coli lacZ* gene. The sequence of the RIPE3a oligonucleotide used was as follows: 5'-GATCTAGCCCCTCTGGCCATCTGCTGATCCG-3'. The E47 yeast expression plasmid containing the *TRP1* gene as a selection marker, YEpE47bHLH, was constructed by inserting a blunt-ended *Ava*I fragment containing the bHLH domain of E47 (nucleotides 1371–2068) (Henthorn et al. 1990) into the *Sma*I site of the vector PCB-Ub stop (Nawaz et al. 1994). The orientation and reading frame were confirmed by sequencing. The protease-deficient strain BJ2168 (*MATa leu2 trp1 ura3-52 prb1-1122 pep4-3 prc1-407 gal2*) was transformed with the YEpE47bHLH and YRpRIPE3a plasmids and selected on plates lacking tryptophan and uracil.

Construction of GAL4 HIT cell cDNA yeast expression library

An oligo(dT)-primed cDNA was generated from HIT cell poly(A)⁺ RNA prepared by standard procedures (Sambrook et al. 1989) using a cDNA synthesis kit (Stratagene) according to the manufacturer's instructions, with one modification. *Nco*I adaptors were prepared and ligated to the 5' end of the cDNA. The sequence of the adaptors are 5'-CATGGCGCAGCTG-3' (upper strand); 5'-CAGCTGCGC-3' (lower strand). The size selected (>500 bp) cDNA was cloned directionally into the *Nco*I-*Xho*I sites of the yeast expression vector pACT(tet)Li, containing the *LEU2* gene as a selection marker. The pACT(tet)Li vector was derived by inserting the tetracycline resistance marker into the *Hpa*I site of pACT (Durfee et al. 1993). A polylinker was inserted into the *Xho*I site of pACT. The sequence of the polylinker used was as follows: 5'-TCGACCCATGGCGGCCGC-CCCCGGGAGCTC-3'. The ligated DNA was transformed into the *E. coli* strain DH10B (BRL) and plated at a density of 1.5×10^4 to 2.0×10^4 colonies/plate on media containing ampicillin. The 500,000 bacterial transformants were scraped from the plates and pooled. Glycerol stocks were prepared by taking 8 × 1-ml aliquots and stored frozen at -70°C . Plasmid DNA was prepared by using MaxiPrep purification Columns (Promega).

Screening the library and β -galactosidase assays

The yeast strain BJ2168 carrying the reporter and expression plasmids were maintained under selection for the *URA3* and *TRP1* markers. These cells were transformed with the unamplified library by the lithium acetate method (Ito et al. 1983). Transformation efficiencies were $\sim 5 \times 10^3$ transformants per microgram of cDNA library. Transformations were plated out

on selective media lacking TRP, URA, and LEU and containing X-gal at a final concentration of 80 mg/l and incubated for 4–5 days at 30°C. Blue colonies were picked and colony purified to obtain single pure positives. These positive colonies were grown overnight at 30°C in 10 ml of selective liquid media. When the cells reached an optical density of 1.0 at 600 nm, they were harvested, lysed by glass bead homogenization in transcription buffer, and the level of β -galactosidase produced was measured as described previously (Miller 1972). The transcriptional buffer for β -galactosidase assays contained 0.12 M Na_2HPO_4 , 0.04 M NaH_2PO_4 , 10 mM KCl, 1 mM MgSO_4 , and 2.7% 2-mercaptoethanol (pH 7.0). The cDNAs from positive cultures were recovered by transforming *E. coli* with plasmid DNA (Ward 1990) then plating on LB/tetracycline plates. The cDNAs were retransformed and reassayed for activity.

Plasmids and sequencing

The *Bgl*III–*Xho*I fragment of BETA2 was subcloned into the *Bam*HI–*Xho*I sites of pGEM7ZF(+). This construct was used for sequencing and in vitro translations. Both strands of BETA2 cDNA were sequenced using Sequenase version 2.0 kit (U.S. Biochemical) according to the manufacturer's instructions. The mammalian expression plasmid for BETA2 was constructed by inserting the *Bgl*III–*Xho*I fragment of BETA2 into the *Bam*HI–*Sma*I sites of pCMV. The in vitro translation plasmid for shPan1 is described elsewhere (German et al. 1991). The E47 and MyoD mammalian expression plasmids are described in Henthorn et al. (1990) and Davis et al. (1987), respectively. The RIPE3 and RIPE3a reporters are described elsewhere (Hwung et al. 1990).

Tissue culture, transfections, and CAT assays

The hamster insulin tumor HIT-T15 M2.2.2 (Santerre et al. 1981) and hamster fibroblast BHK cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, 100 units of penicillin/ml and 100 μ g of streptomycin/ml. Transient transfections were carried out by the calcium phosphate coprecipitation methods as described previously (Edlund et al. 1985; Hwung et al. 1988). Cesium chloride-purified DNA was used for all transfections. For transfections into HIT cells, 1 μ g of reporter and 1 μ g of expression plasmids were used. For transfections into BHK cells, 5 μ g of reporter and 1 μ g of expression plasmids were used. HIT cells were seeded at 1.5×10^6 cells/60-mm dish. BHK cells were seeded at 3×10^5 cells/60-mm dish. Cells were harvested 44–48 hr post-transfection and assayed for CAT activity (Gorman et al. 1982).

Northern blot analysis

RNA was extracted with RNazol B (Cinna/Biotech Laboratories Inc. Houston, TX). Formaldehyde gels were prepared, transferred to Hybond-N membrane (Amersham), and hybridized as described (Stellrecht et al. 1993). Filters were stripped in boiling 0.1% SDS as described by the manufacturer before reprobing.

Gel shift assay

The RIPE3 (–126 to –86) double-stranded oligonucleotide was end-labeled with [α - 32 P]dGTP using the Klenow fragment to $\sim 1.0 \times 10^8$ cpm/ μ g. The RIPE3 sequence used is as follows: 5'-GATCTGGAAACTGCAGCTTCAGCCCCTCTGGCCATCT-GCTGATCCG-3'.

Binding reactions contained 10 mM Tris (pH 7.5), 50 mM NaCl, 1 mM EDTA, 1 mM DTT, 5% glycerol, $\sim 30,000$ cpm of end-labeled oligonucleotide, 3.5 μ g of nuclear extract, or 1 μ l of in vitro-translated protein. Two hundred nanograms of poly[d(I-C)] was used as a nonspecific competitor for in vitro-translated

material and 1.2 μ g was used with nuclear extracts. The mutant E-box oligonucleotides used for competition studies are as follows: 5'-GATCTAGCCCCTCTGGCCAGCTGCTGATCCG-3' (mut 1) and 5'-GATCTAGCCCCTCTGGCCATCTTCTGATCCG-3' (mut 2). After binding, protein–DNA complexes were resolved by electrophoresis on nondenaturing 4% polyacrylamide (19.4:0.6) in $0.5 \times$ TBE (50 mM Tris, 50 mM boric acid, 1 mM EDTA). Protein was synthesized using the Promega TNT-coupled transcription/translation kit according to the manufacturer's instructions. SP6 polymerase was used to synthesize BETA2 and shPan1. Parallel reactions were carried out using [35 S]methionine or unlabeled methionine (1 mM). Labeled protein was resolved on 10% SDS-PAGE to ensure proper synthesis and estimate translation efficiency. Nuclear extracts were prepared as described previously (Peyton et al. 1994).

GST–fusion protein production and purification for antibodies

BETA2 cDNA was digested with *Nco*I, filled in with Klenow, and digested with *Eco*RI. The 1.3-kb fragment was inserted in-frame with the glutathione S-transferase gene of *Sma*I–*Eco*RI-digested pGEX-2T (Pharmacia). The fusion proteins were purified by binding to glutathione–Sepharose and eluting with 100 mM glutathione. The purified material was used as the antigen. Antibodies were raised by Cocalico Biologicals (Reamstown, PA).

Acknowledgments

We thank the members of the Tsai laboratory for critical reading of this manuscript and for helpful suggestions and discussions throughout the course of this work. We are grateful to E. Olson for providing the MyoD–GAL4 yeast expression plasmid, S. Elledge for providing the pACT vector, and D. Hanahan for the STC-1 cell line. This work was supported by a National Institutes of Health grant (HD17379) to M.J.T. and a Juvenile Diabetes Foundation postdoctoral fellowship (FDN 3931121) to C.M.M.S.

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

The nucleotide sequence data reported in this paper have been submitted to the EMBL/GenBank data libraries.

References

- Alpert, S., D. Hanahan, and G. Teitelman. 1988. Hybrid insulin genes reveal a developmental lineage for pancreatic endocrine cell and imply a relationship with neurons. *Cell* **53**: 295–308.
- 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.
- Beckmann, H. and T. Kadesch. 1991. The leucine zipper of TFE3 dictates helix–loop–helix dimerization specificity. *Genes & Dev.* **5**: 1057–1066.
- Begley, C.G., P.D. Aplan, S.M. Denning, B.F. Haynes, T.A. Waldmann, and I.R. Kirsch. 1989. The gene SCL is expressed during early hematopoiesis and encodes a differentiation-related DNA-binding motif. *Proc. Natl. Acad. Sci.* **86**: 10128–10132.
- Benezra, R., R.L. Davis, D. Lockshon, D.L. Turner, and H. Wein-

- traub. 1990. The protein Id: A negative regulator of helix-loop-helix binding proteins. *Cell* **61**: 49–59.
- Caudy, M., H. Vassin, M. Brand, R. Tuma, L.Y. Jan, and Y.N. Jan. 1988. *Daughterless*, a *Drosophila* gene essential for both neurogenesis and sex determination, has sequence similarities to *myc* and the *achaete-scute* complex. *Cell* **55**: 1061–1067.
- Cordle, S.R., E. Henderson, H. Masucha, 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. Mutagenesis of the rat insulin II 5'-flanking region defines sequences important for expression in HIT cells. *Mol. Cell. Biol.* **9**: 1784–1789.
- Dalton, S. and R. Treisman. 1992. Characterization of SAP-1, a protein recruited by serum response factor to the *c-fos* serum response element. *Cell* **68**: 597–612.
- Davis, R.L., H. Weintraub, and A.B. Lassar. 1987. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**: 987–1000.
- Davis, R.L., P.-F. Cheng, A.B. Lassar, and H. Weintraub. 1990. The MyoD DNA binding domain contains a recognition code for muscle-specific gene activation. *Cell* **60**: 733–746.
- Deltour, L., P. Leduque, N. Blume, O. Madsen, P. Dubois, J. Jami, and D. Bucchini. 1993. Differential expression of the two nonallelic proinsulin genes in the developing mouse embryo. *Proc. Natl. Acad. Sci.* **90**: 527–531.
- Durfee, T., K. Becherer, P.-L. Chen, S.-H. Yeh, Y. Yang, A.E. Kilburn, W.-H. Lee, and S. Elledge. 1993. The retinoblastoma protein associates with the protein phosphatase type 1 subunit. *Genes & Dev.* **7**: 555–569.
- 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 elements. *Science* **230**: 912–916.
- Edmondson, D.G. and E.N. Olson. 1993. Helix-loop-helix proteins as regulators of muscle-specific transcription. *J. Biol. Chem.* **268**: 755–758.
- Fields, S. and O.K. Song. 1989. A novel genetic system to detect protein-protein interactions. *Nature* **340**: 245–246.
- German, M.S., M.A. Blonar, C. Nelson, L.G. Moss, and W.J. Rutter. 1991. Two related helix-loop-helix proteins participate in separate cell-specific complexes that bind the insulin enhancer. *Mol. Endocrinol.* **5**: 292–299.
- Gittes, G.K. and W.J. Rutter. 1992. Onset of cell-specific gene expression in the developing mouse pancreas. *Proc. Natl. Acad. Sci.* **89**: 1128–1132.
- Gorman, C.M., L.F. Moffat, and B.M. Howard. 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.* **2**: 1044–1051.
- Guz, Y., M.R. Montminy, R. Stein, J. Leonard, L.W. Gamer, C.V.E. Wright, and G. Teitelman. 1995. Expression of murine STF-1, a putative insulin gene transcription factor, in β -cells of pancreas duodenal epithelium and pancreatic exocrine and endocrine progenitors during ontogeny. *Development* **121**: 11–18.
- Hanahan, D. 1985. Heritable formation of pancreatic β -cell tumors in transgenic mice expressing recombinant insulin/simian virus 40 oncogenes. *Nature* **315**: 115–122.
- Henthorn, P., M. Kiledjian, and T. Kadesch. 1990. Two distinct transcription factors that bind the immunoglobulin enhancer μ E5/ κ E2 motif. *Science* **247**: 467–470.
- Hopwood, N.D., A. Pluck, and J.B. Gurdon. 1989. A *Xenopus* mRNA related to *Drosophila twist* is expressed in response to induction in the mesoderm and the neural crest. *Cell* **59**: 893–903.
- Hsu, J.-S., E. Olson, and R. Kingston. 1992. HEB, a helix-loop-helix protein related to E2A and ITF2 that can modulate the DNA-binding ability of myogenic regulatory factors. *Mol. Cell. Biol.* **12**: 1031–1042.
- Hwung, Y.-P., D.T. Crowe, L.-H. Wang, S.Y. Tsai, and M.-J. Tsai. 1988. The COUP transcription factor binds to an upstream promoter element of the rat insulin II gene. *Mol. Cell. Biol.* **8**: 2070–2077.
- Hwung, Y.-P., Y.-Z. Gu, and M.-J. Tsai. 1990. Cooperativity of sequence elements mediates tissue specificity of the rat insulin II gene. *Mol. Cell. Biol.* **10**: 1784–1788.
- Ito, H., Y. Fukuda, K. Murata, and A. Kimura. 1983. Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol.* **153**: 1632–1638.
- Jan, Y.N. and L.Y. Jan. 1993. HLH proteins, fly neurogenesis, and vertebrate myogenesis. *Cell* **75**: 827–830.
- Jarman, A.P., Y. Grau, L.Y. Jan, and Y.N. Jan. 1993. *atonal* is a proneural gene that directs organ formation in the *Drosophila* peripheral nervous system. *Cell* **73**: 1307–1321.
- Johnson, J.E., S.J. Birren, and D.J. Anderson. 1990. Two rat homologues of *Drosophila achaete-scute* specifically expressed in neuronal precursors. *Nature* **346**: 858–861.
- 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.* **84**: 8819–8823.
- Landschulz, W.H., P.F. Johnson, and S.L. McKnight. 1988. The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. *Science* **240**: 1759–1764.
- Lassar, A.B., R.L. Davis, W.E. Wright, T. Kadesch, C. Murre, A. Voronova, D. Baltimore, and H. Weintraub. 1991. Functional activity of myogenic HLH proteins requires hetero-oligomerization with E12/E47-like protein *in vivo*. *Cell* **66**: 305–315.
- Le Douarin, N.M. 1988. On the origin of the pancreatic endocrine cells. *Cell* **53**: 169–171.
- Lüscher, B. and R.N. Eisenmann. 1990. New light on Myc and Myb: Part I. *Myb*. *Genes & Dev.* **4**: 2025–2035.
- Mellentin, J.D., S.D. Smith, and M.L. Cleary. 1989. *lyl-1*, a novel gene altered by chromosomal translocation in T cell leukemia, codes for a protein with a helix-loop-helix DNA binding motif. *Cell* **58**: 77–83.
- Miller, J.M. 1972. Assays for β -galactosidase. In *Experiments in molecular genetics* (ed. J.M. Miller), pp. 352–355. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Murre, C., P.S. McCaw, and D. Baltimore. 1989a. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, *daughterless*, *MyoD*, and *myc* proteins. *Cell* **56**: 777–783.
- Murre, C., P.S. McCaw, H. Vaessin, M. Caudy, L.Y. Jan, Y.N. Jan, C.V. Cabrera, J.N. Buskin, S.D. Hauschka, A.B. Lassar, H. Weintraub, and D. Baltimore. 1989b. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* **58**: 537–544.
- Nawaz, Z., C. Baniahmad, T.P. Burris, D.J. Stillman, B.W. O'Malley, and M.-J. Tsai. 1994. The yeast SIN3 gene product negatively regulates the transcriptional activity of the human progesterone receptor and positively regulates the transcriptional activity of GAL4 and the HAP1 activator. *Mol. & Gen. Genet.* **245**: 724–733.
- 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.

- Ohlsson, H., O. Karlsson, and T. Edlund. 1988. A beta-cell-specific protein binds to the two major regulatory sequences of the insulin gene enhancer. *Proc. Natl. Acad. Sci.* **85**: 4228–4231.
- Ohlsson, H., S. Thor, and T. Edlund. 1991. Novel insulin promoter- and enhancer-binding proteins that discriminate between pancreatic α and β cells. *Mol. Endocrinol.* **5**: 897–904.
- Olson, E.N. 1990. MyoD family: A paradigm for development? *Genes & Dev.* **4**: 1454–1461.
- Olson, E.N. and W.H. Klein. 1994. bHLH factors in muscle development: Dead lines and commitments, what to leave in and what to leave out. *Genes & Dev.* **8**: 1–8.
- Park, C.W. and M.D. Walker. 1992. Subunit structure of cell-specific E-box-binding proteins analyzed by quantitation of electrophoretic mobility shift. *J. Biol. Chem.* **267**: 15642–15649.
- Pearse, A.G.E. and J. Polak. 1971. Neural crest origin of the endocrine polypeptide (APUD) cells of the gastrointestinal tract and pancreas. *Gut* **12**: 783–788.
- Peyton, M., L. Moss, and M.-J. Tsai. 1994. Two distinct class A helix-loop-helix transcription factors, E2A and BETA1, form separate DNA binding complexes on the insulin gene E-box. *J. Biol. Chem.* **269**: 25936–25941.
- Pham, T.A., J.F. Elliston, Z. Nawaz, D.P. McDonnell, M.-J. Tsai, and B. W. O'Malley. 1991. Antiestrogen can establish nonproductive receptor complexes and alter chromatin structure at target enhancers. *Proc. Natl. Acad. Sci.* **88**: 3125–3129.
- Ptashne, M. 1988. How eukaryotic transcriptional activators work. *Nature* **335**: 683–689.
- Quong, M.W., M.E. Massari, R. Zart, and C. Murre. 1993. A new transcriptional activation motif restricted to a class of helix-loop-helix proteins is functionally conserved in both yeast and mammalian cells. *Mol. Cell. Biol.* **13**: 792–800.
- Robinson, G.L. W.G., M. Peshevaria, E. Henderson, S.-Y. Shieh, M.-J. Tsai, G. Teitelman, and R. Stein. 1994. Expression of the trans-active factors that stimulate insulin control element-mediated activity appear to precede insulin gene transcription. *J. Biol. Chem.* **269**: 2452–2460.
- Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: A laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- 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 beta cells. *Proc. Natl. Acad. Sci.* **78**: 4339–4343.
- Shieh, S.-Y. and M.-J. Tsai. 1991. Cell-specific and ubiquitous factors are responsible for the enhancer activity of the rat insulin II gene. *J. Biol. Chem.* **266**: 16707–16714.
- Staudinger, J., M. Perry, S.J. Elledge, and E. N. Olson. 1993. Interactions among vertebrate helix-loop-helix proteins in yeast using the two-hybrid system. *J. Biol. Chem.* **268**: 4608–4611.
- 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.
- Stellrecht, C.M.M., G. Frazier, C. Selvanayagam, L.-Y. Chao, A. Lee, and G. Saunders. 1993. Transcriptional regulation of a hematopoietic proteoglycan core protein gene during hematopoiesis. *J. Biol. Chem.* **268**: 4078–4084.
- Teitelman, G. and J.K. Lee. 1987. Cell lineage analysis of pancreatic islet cell development: Glucagon and insulin cells arise from catecholaminergic precursors present in the pancreatic duct. *Dev. Biol.* **121**: 454–466.
- Teitelman, G., T.H. Joh, and D.J. Reis. 1981. Transformation of catecholaminergic precursors into glucagon (A) cells in mouse embryonic pancreas. *Proc. Natl. Acad. Sci.* **78**: 5225–5229.
- Teitelman, G., S. Alpert, J.M. Polak, A. Martinez, and D. Hahan. 1993. Precursor cells of mouse endocrine pancreas co-express insulin, glucagon and the neuronal proteins tyrosine hydroxylase and neuropeptide Y, but not pancreatic polypeptide. *Development* **118**: 1031–1039.
- Thisse, B., C. Stoetzel, C. Gorostiza-Thisse, and F. Perrin-Schmitt. 1988. Sequence of the *twist* gene and nuclear localization of its protein in endomesodermal cells of early *Drosophila* embryos. *EMBO J.* **7**: 2175–2183.
- Villares, R. and C.V. Cabrera. 1987. The achaete-scute gene complex of *D. melanogaster*: Conserved domains in a subset of gene required for neurogenesis and their homology to *myc*. *Cell* **50**: 415–424.
- Voronova, A. and D. Baltimore. 1990. Mutations that disrupt DNA binding and dimer formation in the E47 helix-loop-helix protein map to distinct domains. *Proc. Natl. Acad. Sci.* **87**: 4722–4726.
- Walker, M.D., T. Ecklund, A.M. Boulet, and W.J. Rutter. 1983. Cell-specific expression controlled by the 5'-flanking region of insulin and chymotrypsin genes. *Nature* **306**: 557–561.
- 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.
- Ward, A.C. 1990. Single-step purification of shuttle vectors from yeast for high frequency back-transformation into *E. coli*. *Nucleic Acids Res.* **18**: 5319.
- Weintraub, H. 1993. The MyoD family and myogenesis: Redundancy, networks, and thresholds. *Cell* **75**: 1241–1244.
- Weintraub, H., R. Davis, S. Tapscott, M. Thayer, M. Krause, R. Benzra, T. K. Blackwell, D. Turner, R. Rupp, S. Hollenberg, Y. Zhuang, and A. Lassar. 1991. The MyoD gene family: Nodal point during specification of the muscle cell lineage. *Science* **251**: 761–766.
- Weintraub, H., T. Genetta, and T. Kadesch. 1994. Tissue-specific gene activation by MyoD: Determination by cis-acting repression elements. *Genes & Dev.* **8**: 2203–2211.
- Whelan, J., D. Poon, P. A. Weil, and R. Stein. 1989. Pancreatic β -cell-type-specific expression of the rat insulin II gene is controlled by positive and negative transcriptional elements. *Mol. Cell. Biol.* **9**: 3253–3259.
- Whelan, J., S.R. Cordle, E. Henderson, P.A. Weil, and R. Stein. 1990. Identification of a pancreatic β -cell insulin gene transcription factor that binds to and appears to activate cell-type-specific expression: Its possible relationship to other cellular factors that bind to a common insulin gene sequence. *Mol. Cell. Biol.* **10**: 1564–1572.
- Wolf, C., C. Thisse, C. Stoetzel, B. Thisse, P. Gerlinger, and F. Perrin-Schmitt. 1991. The *m-twist* gene of *Mus* is expressed in subsets of mesodermal cells and is closely related to the *Xenopus x-twist* and the *Drosophila twist* genes. *Dev. Biol.* **143**: 363–373.
- Xia, Y., L. Brown, C.Y.-C. Yang, J.T. Tsan, M.J. Siciliano, R. Espinosa III, M.M. LeBeau, and R.J. Baer. 1991. TAL2, a helix-loop-helix gene activated by the (7;9)(q34;q32) translocation in human T-cell leukemia. *Proc. Natl. Acad. Sci.* **88**: 11416–11420.



Tissue-specific regulation of the insulin gene by a novel basic helix-loop-helix transcription factor.

F J Naya, C M Stellrecht and M J Tsai

Genes Dev. 1995, **9**:

Access the most recent version at doi:[10.1101/gad.9.8.1009](https://doi.org/10.1101/gad.9.8.1009)

References

This article cites 72 articles, 37 of which can be accessed free at:
<http://genesdev.cshlp.org/content/9/8/1009.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](#).

Dharmacon[™] Reagents
Custom synthesis, RNAi, and CRISPR solutions

Infinite
Reliability

More

horizon
a PerkinElmer company