J. Biochem. 126, 494-502 (1999)

Human Subtilisin-Like Proprotein Convertase, PACE4 (SPC4) Gene Expression Is Highly Regulated through E-Box Elements in HepG2 and GH4C1 Cells[1]

Akihiko Tsuji, Shigeru Yoshida,[2] Shin-ichi Hasegawa, Miwa Bando, Ichiro Yoshida, Shizuyo Koide, Kenji Mori, and Yoshiko Matsuda[3]

Department of Biological Science and Technology, Faculty of Engineering, The University of Tokushima, 2-1 Minamijosanjima, Tokushima 770-8506

Received April 5, 1999; accepted June 18, 1999

PACE4 (SPC4) is a member of the mammalian subtilisin-like proprotein convertase (SPC) family, which participates in maturation of precursor proteins. PACE4 is expressed at high levels in the anterior pituitary, central nervous system, the developing olfactory bulb, heart, and liver. Recently, we determined the gene structure of human PACE4. [Tsuji *et al.* (1997) *J. Biochem.* 122, 438-452]. The 5′-flanking region of PACE4 gene contains 12 E-boxes (E1 to E12) within 1 kb upstream of the transcription initiation site. To examine the function of these E-box elements in the regulation of PACE4 expression, deletion and mutation constructs of the 5′-flanking region were ligated to the luciferase gene and analyzed for promoter activity in HepG2 and GH4C1 cells, which express PACE4 at high level. Some differences were observed in the activity of each promoter construct between HepG2 and GH4C1 cells, although the overall profiles of activity for the promoter fragment series were similar regardless of cell type. We showed that the basal promoter activity of the PACE4 gene is first determined by sequences lying between −315 and −1 bp and further regulated by positive and negative elements in the upstream region. Site-directed mutagenesis of E-boxes in these regulatory elements showed that the E10 E-box act as positive regulator, whereas an E-box cluster (E4-E9) acts as a negative regulator in both cells. E2 E-box acts as a positive regulator only in HepG2 cells. Other E-boxes (E1, E3, and E12) had no effect on the promoter activity. These results indicate that E-box elements play a critical role in controlling PACE4 expression in HepG2 and GH4C1 cells and that PACE4 expression is regulated by a mechanism distinct from that of other SPC family proteases.

Key words: bHLH, E-box, PACE4, proprotein convertase, transcriptional regulation.

In mammals, seven members of the subtilisin-like proprotein convertase (SPC) family, which are involved in the proteolytic processing of precursor proteins at sites consisting of paired basic amino acid residues, have been identified to date (*1, 2*). These include furin/SPC1 (*2*), PC2/SPC2 (*3*), PC1/3/SPC3 (*4*), PACE4/SPC4 (*5-7*), PC4/SPC5 (*8*), PC5/6/SPC6 (*9*), and LPC/PC8/SPC7 (*10-12*). These enzymes are Ca^{2+}-dependent serine proteases. All the SPCs share the structural similarities of a signal peptide, propeptide, subtilisin-like catalytic domain, and homoB domain as the basic domain structure. The carboxy terminal region varies greatly in length and sequence among SPCs. Tissue distribution and intracellular localization of these SPCs are also different. Furin and LPC/PC8, which have their transmembrane region at the carboxy terminus, are expressed in a wide variety of tissues (*1, 2, 10, 11*). PC2, PC1/3 and PC4 are predominantly expressed in endocrine tissues (*3, 4, 8, 13*). PACE4 and PC5/6 show restricted distribution in both endocrine and nonendocrine tissues, but their profiles of tissue distribution differ (*1, 13-15*). PACE4 is found at high levels in the anterior pituitary, cerebellar Purkinje cells, B cells of pancreatic islets, mitral cells of the olfactory bulb, heart, and liver. Recently, we showed spatio-temporal expression of PACE4 in the rat olfactory bulb (*15*). Moreover, Constam *et al.* reported that mouse PACE4 exhibits highly regulated expression patterns during embryogenesis (*12*). These results suggest a unique physiological function of PACE4. Constam and Robertson also showed that PACE4 can process proBMP (bone morphogenetic protein) 4 to active BMP4 by cotransfection experiments (*16*). More recently, we demonstrated that PACE4, as well as furin and LPC/PC8, is involved in the processing of proalbumin in HepG2 cells by coexpression and antisense-RNA studies (*17*).

Understanding the mechanism governing tissue-specific expression of the PACE4 gene requires knowledge of the promoter function of the gene. In our effort to elucidate the regulatory mechanism of human PACE4 expression, we cloned and studied the structure of the human PACE4 gene

[1] This work was supported by Grants-in-Aid for Scientific Research (C) and Priority Area (Intracellular Proteolysis) from the Ministry of Education, Science, Sports and Culture of Japan.

[2] Present address: Pharmacology Laboratories, Yamanouchi Pharmaceutical Co. Ltd., 21-Miyukigaoka, Tsukuba, Ibaraki 305-0841.

[3] To whom correspondence should be addressed. Tel: +81-88-656-7523, Fax: +81-88-655-3161, E-mail: matsuda@bio.tokushima-u.ac.jp

(*18*). The PACE4 gene spans at least 250 kb and is distributed over 25 exons. Compared with other SPCs such as furin (*19, 20*), PC1/3 (*21*), PC2 (*22*), PC4 (*23*), and PC8 (*24*), PACE4 has the most complicated SPC gene structure reported to date. Sequence analysis of the 5′-flanking region revealed that the PACE4 gene lacks TATA and CCAAT boxes in the proximal upstream region of the transcription start site. Potential binding sites for several transcription factors including SP1, AP1, AP2, PEA3, Ets-1, GHF-1, CREB, and basic helix-loop-helix (bHLH) proteins were found in this region. The most interesting feature of the sequence of this region is the presence of 12 E boxes (CANNTG, E1-E12): the target sequence for bHLH transcription factors. The bHLH factors are developmental regulators of transcription in a variety of tissues including muscle, nervous tissues, and blood cell lineages (*25-27*). Many mammalian genes with E boxes in their control elements show cell-type specific expression and are probably regulated by bHLH transcription factors. The fact that these E boxes are conserved in terms of position and sequence between rat and human PACE4 promoter regions strengthens the idea of their potential role in specific spatial and temporal expressions of PACE4.

In this study, we addressed the question of whether the E box motif is involved in the regulation of PACE4 expression in HepG2 and GH4C1 cells using deletion analysis and site-directed mutagenesis in transient transfection experiments. GH4C1 cells, a rat pituitary somatomammotroph tumor cell line, expressed PACE4 at high levels, although its physiological role remains to be clarified. Our results showed that the E-box cluster (E4-E9) was a strong negative acting element, whereas the E10 E-box was a positive element in both HepG2 and GH4C1 cells. On the other hand, the E2 E-box was a positive element in HepG2 cells not in GH4C1 cells.

EXPERIMENTAL PROCEDURES

Northern Blot Analysis—Total RNAs were isolated from various culture cells using ISOGEN (Nippon Gene, Toyama) according to the manufacturer's protocol. Five micrograms of each total RNA were resolved by electrophoresis in an agarose gel containing 6.7% formaldehyde and transferred onto a nylon membrane (Hybond-N+, Amersham Pharmacia Biotech, Buckinghamshire, England). Hybridization was carried out using a ^{32}P-labeled PACE4-homoB cDNA probe as described previously (*15*).

Construction of Reporter Plasmid—Human PACE4 promoter regions of various lengths and their point mutants were prepared by PCR amplification (*28*) because of the lack of suitable restriction sites within this upstream region of the PACE4 gene. The 3′-sites of all PCR products included sequences up to but not including the ATG initial codon, except for the −982/−315 construct. Human genomic DNA fragment (about 8 kb) isolated from λ phage clone λhPA OL-3 by *Bam*HI and *Hin*dIII digestion was used as the PCR template. This fragment contained 2.5 kb of 5′-flanking region, 0.6 kb of exon 1, and 5 kb of intron A. PCR was carried out in the presence of dimethyl sulfoxide using *Xho*I-linked sense primer and *Hin*dIII-linked antisense primer. The internal deletion mutant −982 (del −780/−649) was prepared by PCR as follows. The 3′-terminal fragment lacking the E-box cluster was generated by PCR using 5′-CGAAGCGAGGTTAAGACGGCCTGCCTCCCA-3′ (−797 to −783 and −649 to −635) as the sense mutagenic primer and 5′-CCCAAGCTTAGCGGCGACAGGCTCGCGCG-3′ (−20 to −1, P1 primer) as the antisense primer. The amplified fragment was ligated to the 5′-terminal fragment (−982 to −798) and used as a template for the second PCR. The internal deletion mutant was generated by the second PCR using −982 sense primer (5′-CCGCTCGAGTACATCGTCTAGTGAACACC-3′) and P1 antisense primer. Amplified fragments were subcloned into *Xho*I and *Hin*dIII sites of pGL3-Basic vector (Promega, Madison, WI, USA). The sequence of the insert was confirmed.

Cell Culture and Transfection—HepG2 and GH4C1 cells were grown at 37°C under 5% CO_2 in low glucose (1 g/liter) and high glucose (4.5 g/liter) Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum and antibiotics, respectively. The lipofection method using Trans IT™-LT1 transfection reagent (Mirus, Madison, WI, USA) was used for transfection of cells with reporter plasmid according to the manufacturer's instructions. Trans IT reagent (4 μl) was incubated in 100 μl of opti-MEM (Gibco-BRL, Grand Island, NY, USA) at room temperature. After 5 min, reporter plasmid (1.3 μg) and pCMV-β-galactosidase (0.7 μg) were added, and the mixture was incubated for 5 min at room temperature. The cells (50% confluent) in a 60-mm dish which had been seeded 24 h previously were washed with phosphate-buffered saline, and 2 ml of opti-MEM was added. Then TransIT-plasmid mixture (100 μl) was added, and the cells were incubated at 37°C for 4 h. The medium was then replaced with fresh DMEM-10% fetal calf serum, and incubation was continued for 2 days. Cell lysis and luciferase assays were conducted using Promega's luciferase assay system. Luciferase measurements were performed on a Lumat LB9501 luminometer (Berthold, Germany). β-Galactosidase activity was assayed using *o*-nitrophenyl-β-D-galactopyranoside (*29*). The luciferase activities were normalized on the basis of β-galactosidase activity. The results are presented as the means of the luciferase activities±SD from a single experiment performed in triplicate. The experiments were repeated 5-6 times.

Preparation of Nuclear Extract and Gel Mobility Shift Assay—Nuclear extracts were prepared from HepG2 cells and GH4C1 cells as described by Dignam *et al.* (*30*) with a slight modification (*29*). The nuclear extract was stored at −80°C until use. Protein concentration was determined by the method of Bradford with bovine serum albumin as the standard (*31*). The sequences of the top strands of oligonucleotides used as probes or competitors in gel mobility shift assay were as follows: wild-type E2 E-box, 5′-AGGGTCTGAGGCCCCCGCCAGGTGAGTGCG-3′ (−420 to −391); mutated E2 E-box (EM2), 5′-AGGGTCTGAGGCCCCCGC**TAGGTC**AGTGCG-3′; wild-type E9 E-box, 5′-GGCCTGGGGGTTCACCTGCACGGCCTGGGGGTT-3′ (−773 to −741); mutated E9 E-box (EM9), 5′-GGCCTGGGGGTT**GTCCTG**CACGGCCTGGGGGTT-3′; wild-type E10 E-box, 5′-CCACCACAGTCATTTGCCCAAGATCG-3′ (−823 to −792); mutated E10 E-box, 5′-CCACCACAGT**GATTTC**CCCAAGATCG-3′; and standard E-box, 5′-GATCCCCCCAACAGGTGCTGCCTGA-3′. The wild-type E-box sequence is underlined, and the mutated E-box sequence is in boldface. SP1 oligonucleotide (5′-ATTCGA-

TCGGGGCGGGGCGAGC-3′) and AP-2 oligonucleotide (5′-GATCGAACTGACCGCCCGCCCGT-3′) were purchased from Promega. The complementary strand of each oligonucleotide was annealed and purified by polyacrylamide gel electrophoresis, and the double-stranded probe was labeled with [γ-^{32}P]ATP using polynucleotide kinase. The DNA binding reaction was performed in 20 μl of reaction mixture containing 0.25 μg of poly(dI-dC)-poly(dI-dC), 10 mM Tris-HCl, pH 7.5, 1 mM $MgCl_2$, 50 mM NaCl, 0.5 mM EDTA, 0.5 mM dithiothreitol, 4% glycerol, 5 μg of cell nuclear extract, and radiolabeled probe (5×10^4 cpm). The mixture was incubated at room temperature for 30 min and loaded on a 4% polyacrylamide gel. Electrophoresis was performed at a constant 150 V for 2 h. The gel was dried and analyzed by use of a BAS-1500 bioimaging analyzer (Fuji Photo Film, Tokyo).

RESULTS AND DISCUSSIONS

Cell Specific Expression of PACE4—To identify cell lines that could be used as models of tissue-specific regulation of the PACE4 gene, we first analyzed PACE4 mRNA levels in various cell lines by Northern blotting. As shown in Fig. 1, PACE4 transcript (4.4 kb) was identified in AtT-20, GH4C1 (rat pituitary cell lines), IMR-32 (human neuroblastoma), HepG2 (human hepatocellular carcinoma), A-549 (human lung carcinoma), and MIA-PaCa (human pancreatic carcinoma) cells. HeLa (human cervical carcinoma) and PC-12 (rat adrenal phenochromonocytoma) cells did not express PACE4. When human PACE4-homoB cDNA probe was used, the PACE4 transcript seemed to be most abundant in HepG2 cells (Fig. 1, left). However, a higher level of PACE4 expression was observed in GH4C1 cells when rat homoB cDNA probe was used (Fig. 1, right). PACE4 transcript was not detected in JM (human T-cell lymphocyte) or NALM6 (human B-cell lymphocyte) (data not shown). Thus HepG2 and GH4C1 cells expressed PACE4 at the highest level, and we selected these cell lines for the promoter analysis. Previously we showed that a major transcription start site is 314 bp upstream from the ATG translation start site in HepG2 cells by primer extension analysis (*18*). Other sites, visible only by overexposing the gel, are located at −139 and −130 bp. The transcription start in GH4C1 cells was also analyzed by the same method. Primer extension analysis using total RNA from GH4C1 cells resulted in a single band (data not shown). The site is located 152 bp upstream from the translation initiation codon.

Mapping of the Promoter Region of the PACE4 Gene—Within the 1-kb 5′-upstream region of human PACE4 gene, various potential binding sites for the transcription factors were identified by searching the transcription factor database (*18*). Twelve putative E-boxes (−1032/−1027, −839/−834, −807/−802, −761/−756, −740/−735, −719/−714, −698/−693, −677/−673, −656/−651, −558/−553, −402/−397, and −352/−347), which are binding sites for the developmentally regulated and cell type-specific bHLH transcription factors, are located in this region. These E boxes were designated E1 to E12 in the order of their distance from the translation start site of PACE4, as shown in Fig. 2 (top). Four more E-boxes (E13, −1348/−1343; E14, −1369/−1364; E15, −1483/−1478; and E16, −1503/−1498) were located further upstream. There are also potential binding sites for SP1 (−436/−431, −407/−402, −292/−287, −259/−254, −173/−166, and −133/−127), AP2 (−954/−947, −829/−822, −618/−609, −456/−449, −431/−424, −409/−403, −84/−77, and −48/−38), CREB (−895/−888), and IL6-responsible element (−474/−469). To investigate the functional activities of E-boxes and other elements, we first constructed various PACE4 5′-deletion mutants and an internal deletion mutant that were fused to the promoterless luciferase reporter gene (Figs. 2 and 3). Each vector was transfected into HepG2, GH4C1, and HeLa cells by the lipofection method. As shown in Fig. 2 (bottom), the profiles of luciferase activity of deletion mutants in HepG2 and GH4C1 cell lines are basically similar, although some differences are observed. In HeLa cells, which do not express PACE4 mRNA as shown in Fig. 1, −395, −796, −982, and −1500 constructs had no significant activity (Fig. 2). These results suggest that the element capable of directing cell-specific expression of the PACE4 gene is located in the upstream region. There are two major regions of positive regulatory activity in HepG2 and GH4C1 cell lines. The promoter activity was completely lost by deletion of the region between −315 and −1 (−982/−315 construct), whereas the −160 construct still showed 68 and 91% of the maximum promoter activity of the −649 construct in HepG2 and GH4C1 cells, respectively. These results indicated that the element located in the region between −315 and −1 is essential for basal expression of PACE4 in both cell lines. There are two potential binding sites for SP1 and AP-2 in the region between −160 and −1. To determine the functional elements in the region between −160 and −1 in more detail, two deletion mutants, −77 and −47, were constructed and their promoter activities were analyzed. The former construct contained two AP-2 sites, whereas the latter contained one AP-2 site. The −77 and −47 constructs showed about 60 and 20% of the promoter activity of the −160 construct in HepG2 and GH4C1 cells, respectively. Thus these results suggested that the potential binding sequences for SP1 (−133/−127) and AP-2 (−84/−77), which are located at downstream of the transcription start site, are important for the basal promoter activity of the PACE4 gene in both HepG2 and GH4C1 cells. Several reports have described transcriptionally active sequences downstream of the transcriptional start sites in various genes (*32, 33*). Miyagoe *et al.* identified promoter elements downstream of the cap site in mouse complement C4 gene (*32*). They showed that these elements play an important role in both the precise determination of the transcriptional initiation site and basal transcription of the gene. It is highly likely that the crucial element for basal and accurate transcription of the PACE4

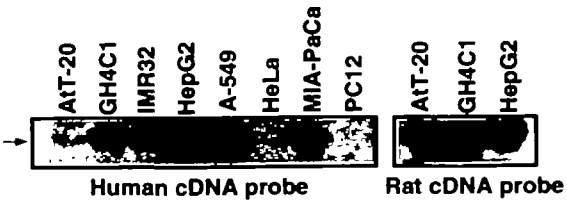

Fig. 1. **Northern blot analysis of PACE4 gene expression in various culture cells.**

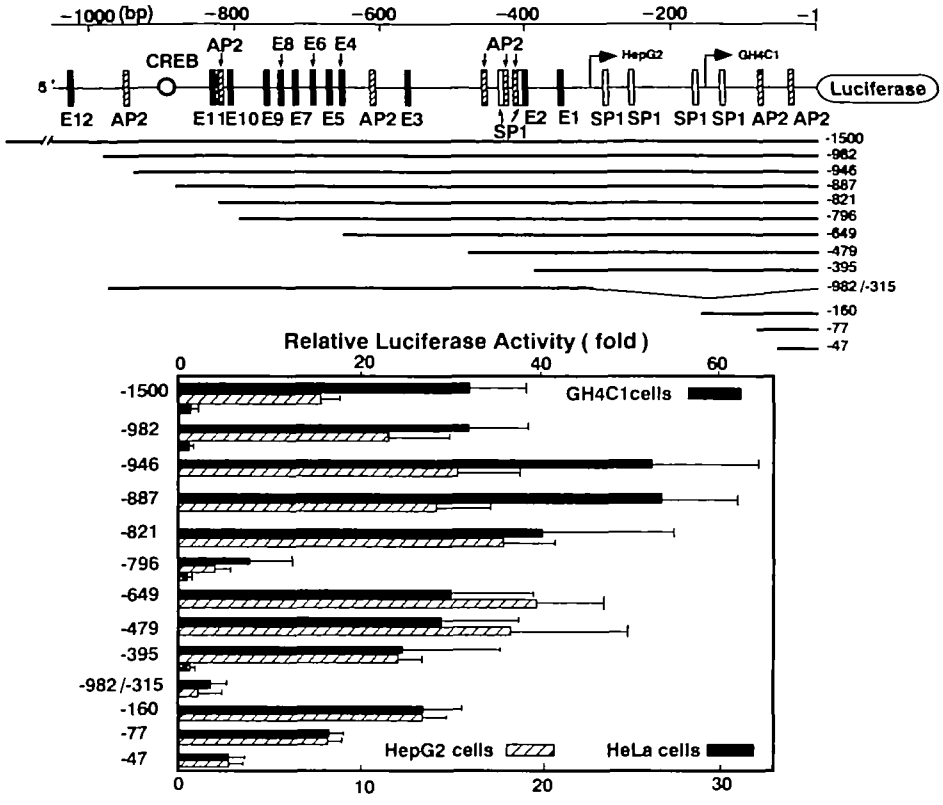

Fig. 2. **Construction of luciferase expression vectors and transfection analysis of human PACE4 gene promoter in HepG2, GH4C1, and HeLa cells.** A series of 5′-end deletion mutants of the promoter fragment were ligated into a luciferase plasmid (pGL3-Basic). The construct −982/−315 contains the internal deletion (positions −314 to −1). Numbers indicate base pairs from the start of translation. Putative regulatory elements are also shown at the top of the figure (hatched box, E-box; dashed box, AP2; open box, SP1; open circle, cAMP response element). E-boxes are numbered from E1 to E12 according to their distance from the translation start site. The sequence of E-boxes is as follows: E1, CACCTG; E2, CAGGTG; E3, CACATG; E4-E9, CACCTG; E10, CATTTG; E11, CAGGTG; E12, CAGCTG. The transcription start sites, −314 in HepG2 cells and −152 in GH4C1 cells, are indicated by bent arrows. The luciferase activity was normalized with the β-galactosidase. The luciferase activity of each construct is expressed relative to the base line luciferase activity of a promoterless construct (pGL3-Basic vector) set to 1.

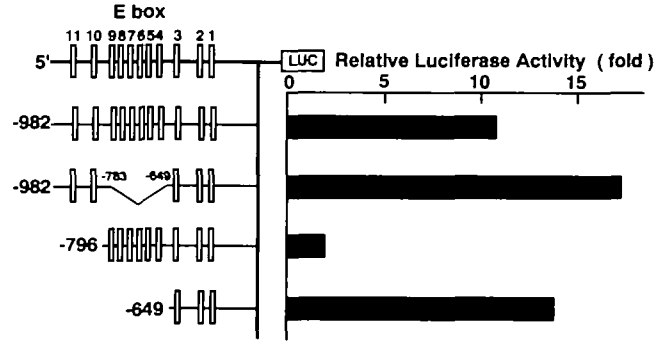

Fig. 3. **Effect of internal deletion of E-box cluster (E4-E9) on the promoter activity in HepG2 cells.** A series of 5′-deletion mutants and internal deletion mutant −982 (del −783/−649) were transfected into HepG2 cells. The luciferase activity of each construct is expressed relative to the luciferase activity of a promoterless pGL3-Basic construct on the right. The results presented are the average from at least three experiments performed in triplicate.

gene is located in the region between −315 and −1. Different transcriptional initiation sites might be used by the −160, −77, and −47 constructs due to deletion of this element.

On the other hand, various negative and positive regulatory elements were found in the upstream region. We observed that the region from −796 to −649 contains a strong silencer activity in both cell lines. Deletion of this region caused 11- and 4-fold increase in luciferase activity in HepG2 and GH4C1 cells, respectively. To confirm silencer activity in this region, the internal deletion mutant was constructed and transfected into HepG2 cells. Compared with the −982 construct, the internal deletion of this region (−782 to −650) caused a 1.6-fold increase in luciferase activity (Fig. 3). The sequence from −982 to −946 also contains a weak silencer activity. Deletion of this sequence caused 1.3- and 1.6-fold increases in the activity in HepG2 and GH4C1 cells, respectively. In contrast, there is a major positive region between −821 and −796 (10-fold change in HepG2 cells, 5-fold change in GH4C1 cells). We also showed a positive regulatory region between −479 and −395 (1.6-fold change in activity) in HepG2 cells. The positive element was located between −887 and −821 in GH4C1 cells (1.3-fold change in activity). These results thus highlighted three features of PACE4 gene regulation in HepG2 and GH4C1 cells. First, the basal promoter activity of the PACE4 gene is determined by the element(s) lying between −160 and −1 bp downstream of the transcription initiation site. Second, the region −796 to −649 appeared to be important for negative regulation. Third, the region (−821 to −796) suppresses the silencer activity of the element located in the region from −796 to −649.

There are possible CRE (TGACCTCA, −895/−888) and IL6-responsible (CTGGGA, −474/−469) elements in the upstream region of the PACE4 gene. Expression of PC1/3, which is involved in the tissue-specific processing of prohormones and neuropeptide precursors within the regulated secretory pathway in (neuro)endocrine cells, is controlled by cAMP (*34, 35*). Effects of cAMP on the promoter activity were also determined in HepG2 and GH4C1 cells using deletion mutants. However, the promoter activity was not changed, suggesting that the CREB element is not functional in either cell. IL6 also had no effect on the promoter activity of the PACE4 gene in both cells.

Effects of Mutagenesis at the E-Boxes on the Promoter Activity and Gel Shift Mobility Assay—The results in Fig.

2 indicated that the upstream region of the PACE4 gene contains several positive and negative regulatory sequences. Several of the 12 E-boxes are located in the positive (E2 and E10) and negative sequences (E4-E9 E-box cluster). Next we determined the activity of these E-boxes by mutation analysis. The positive sequence between −479 and −395 contained one E-box (E2), two SP1 and three AP-2 sites as shown in Fig. 4 (left). To determine which of these response elements is functional in the expression of the PACE4 gene, these elements in the −479 construct were modified by site-directed mutagenesis and their promoter activities were determined. As shown in Fig. 4, mutation at E1 E-box, SP1 and AP2 sites had no significant effect on the promoter activity in HepG2 and GH4C1 cells. However, the mutation of E2 E-box caused a 70% decrease in luciferase activity in HepG2 cells. In contrast, mutation at E2 E-box had no effect on the promoter activity in GH4C1 cells. These results indicate that the E2 E-box functions as a positive regulatory element in HepG2 cells. To identify the binding activity of the endogeneous E-box binding protein that interacts with the E2 E-box, we performed a gel mobility shift assay using nuclear proteins from HepG2 cells and an oligonucleotide (E2 probe) corresponding to −420 to −391. This oligonucleotide contains SP1, AP2, and E2 E-box motifs. Incubation of the E2 probe with HepG2 nuclear extract produced three major retarded bands (Fig. 5). It was most likely that the upper band was specific E2-protein complex. The upper band competed with both an excess amount of unlabeled E2 probe and standard E-box oligonucleotide, but did not compete with SP1 and AP2 oligonucleotide. Furthermore, when an excess of unlabeled competitor containing a mutation in the E2 E-box sequence (CAGGTG in E2 probe was mutated to TAGGTC) was added to the binding reaction, the upper band did not compete. This specific band did not shift to the top when anti-SP1 and anti-AP2 antibodies were added after the binding reaction (data not shown). The middle and lower bands barely competed with excess amount of standard E-box. The lower band competed with mutated E2 E-box oligonucleotide. These results indicate the presence of the specific E2-binding protein in the HepG2 nuclear extract.

As mentioned above, the sequence between −796 and −649 acts as a strong negative element in both HepG2 and GH4C1 cells. In particular, six repeats of GGCCTGGGGT-TCACCTGC sequences containing an E-box (underlined) are located in this region. To investigate the activity of these E-boxes, two mutants were constructed. As shown in Fig. 6, mutation of E4 or E9 E-boxes among the six E-boxes caused a 4.2- or 3.9-fold increase in activity compared with wild-type −796-construct in HepG2 cells. These activities represented almost half the activity of the −649 construct, which contained no E-box cluster. Similar results were obtained in GH4C1 cells (data not shown). These effects indicate that E-boxes between E4 and E9 have a strong negative effect on the promoter activity of the PACE4 gene. Further analysis is necessary to clarify the function of E-boxes E5-E8. To identify the endogenous E9 E-box binding protein, we performed a gel mobility shift assay using nuclear proteins from HepG2 cells and an oligonucleotide (E9 probe) corresponding to −773 to −741 (5′-GGCCTGGGGGTTCACCTGCACGGCCTGGGGGTT-3′). Incubation of the E9 probe with HepG2 nuclear extract produced a single retarded band (data not shown).

The sequence (−821 to −796) identified as a positive element by deletion analysis contains an E-box (E10). The function of the E10 E-box element as a positive regulator

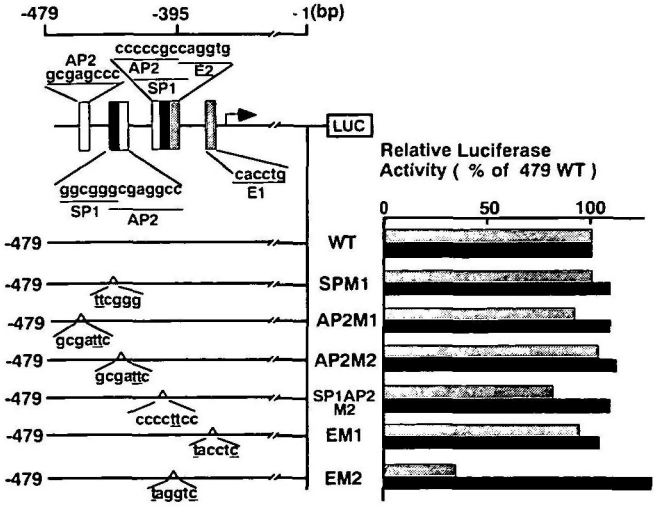

Fig. 4. **Effects of site-directed mutagenesis of potential regulatory elements located within 479 bp upstream of the PACE4 promoter.** The sequence of possible binding sites for transcription factors such as E-boxes (E1 and E2), AP2 and SP1 in the 5-flanking region were mutated as shown in the figure (mutated sequence indicated by underline). A luciferase construct was transfected into the HepG2 and GH4C1 cells. The luciferase activity (HepG2 cells, hatched box; GH4C1 cells, closed box) was expressed as a percentage of the activity of the wild-type construct (top right).

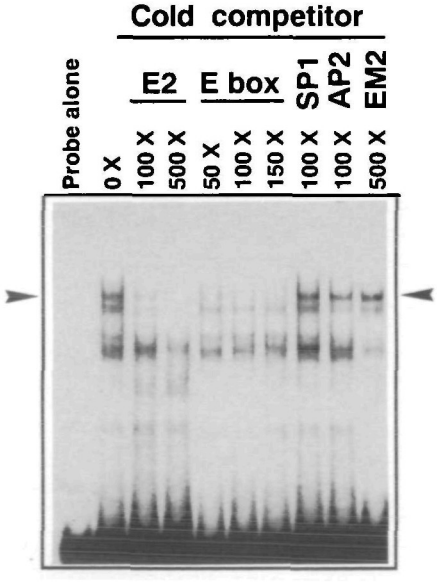

Fig. 5. **Gel mobility shift analysis of nuclear proteins interacting with the E2 E-box sequence.** The ^{32}P-labeled double strand E2 E-box oligonucleotide (E2 probe) incubated with nuclear proteins (5 μg) from HepG2 cells in the presence and absence of cold excess E2-probe (100 and 500-fold molar excess), standard E-box (50, 100, and 150-fold molar excess), SP1 (100-fold), AP2 (100-fold), and EM2 oligonucleotides (500-fold) as described under "EXPERIMENTAL PROCEDURES."

was confirmed by site-directed mutagenesis. The −821-EM10 construct (CATTTG was mutated to TATTTC) exhibited about 42% activity compared to the −821 construct, as shown in Fig. 6. These results indicate that E10 acts as a positive regulator. In addition, the mutation of E11 caused a slight decrease of luciferase activity, suggesting that E11 is a weak positive regulator (data not shown). To identify the binding activity of endogeneous E10 E-box binding protein, a gel mobility shift assay using an oligonucleotide (E10 probe) corresponding to −823 to −792 was performed. Incubation with E10 probe and HepG2 nuclear protein produced a single retarded band, as shown in Fig. 7. This band competed fully with excess unlabeled probe and partially with excess unlabeled standard E-box oligonucleotide. However, it did not compete with mutated E10 E-box oligonucleotide (EM10). All bHLH proteins bind specifically to a consensus hexametric DNA sequence, E-box (CANNTG). E-box sequences are often palindromic and contain identical half-sites, each of which may be bound by one component of the dimer. Although the central dinucleotides of the E-box are usually GC or CG, exceptions have been noted. Additional sequence specificities flanking the consensus E-box element are also reported. The central dinucleotide sequences of E-box in E10 probe (5′-CCACCACAGT**CATTTG**CCCAAGATCG-3′) and standard E-box oligonucleotide (5′-GATCCCCCCAA**CAGGTG**CTGCCTGA-3′) are TT and GG, respectively. It was highly likely that the difference in these central dinucleotide sequences in E-box elements caused partial competition with standard E-box oligonucleotide. These findings suggest that E10 in the segment participates mainly in the positive regulation of the PACE4 gene. However, participation of other elements located in this region is not excluded, although the region between −821 and −796 has no other known response elements.

Recently, the developmental regulation of PACE4 was reported by Constam *et al.* (*12*) and ourselves (*15*). The temporary and spatially tightly-regulated expression pattern during embryogenesis is consistent with a possible involvement of PACE4 in the control of tissue differentiation. Moreover, PACE4 exhibited restricted unique tissue distribution. It is expressed in nonendocrine tissues such as the liver, and endocrine tissues such as the anterior pituitary, heart, and pancreatic islets. However, the molecular mechanisms responsible for the spatial, temporal, and cell-specific expression of PACE4 are still poorly understood. In this study, we showed that the sequence located between −315 and −1 bp downstream of the transcription initiation site is essential for basal expression of PACE4, and its expression is further regulated positively and negatively through E-box elements. Thus, transcription of the PACE4 gene occurs under complex regulation, and the sequences responsible for regulating the tissue-specific expression of PACE4 gene remain to be clarified. The profiles of promoter activity of the deletion constructs were broadly similar between HepG2 and GH4C1 cells, but some differences were also observed. The positive regulatory activity of E2 E-box in HepG2 cells was not observed in GH4C1 cells. As shown in Fig. 1, IMR-32 (human neuroblastoma), A-549 (human lung carcinoma), and MIA-PaCa (human pancreatic carcinoma) cells also express PACE4. Comparative analysis of the transcriptional regulation in these different cell lineages may provide further clues for understanding the tissue-specific expression of PACE4 gene. We have demonstrated the importance of E-box elements in the regulation of the PACE4 gene. The

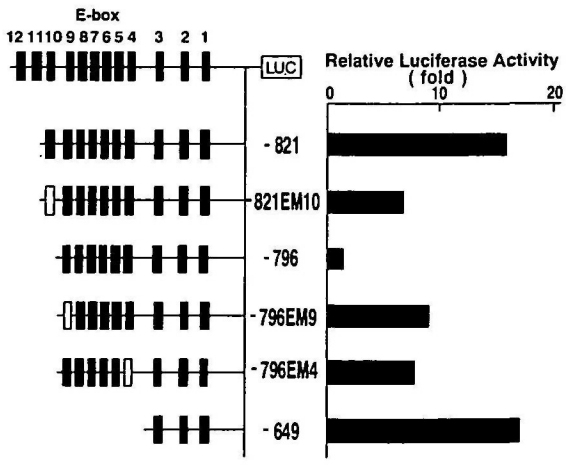

Fig. 6. **Effects of site-directed mutagenesis of E-box consensus sequence (E4, E9, and E10) on the promoter activity.** A series of E-box mutants of the PACE4 gene promoter were generated in the luciferase vector and transfected into HepG2 cells. E-box sequence of E4, E9 (CACCTG), E10 (CATTTG), and E11 (CAGGT) are mutated to GTCCTG, TATTTC, and TAGGTC (mutated sequences are underlined), respectively. Wild type and mutant of E-box are indicated by closed and open squares. The luciferase activity of each construct is expressed relative to the luciferase activity of a promoterless pGL3-Basic construct on the right.

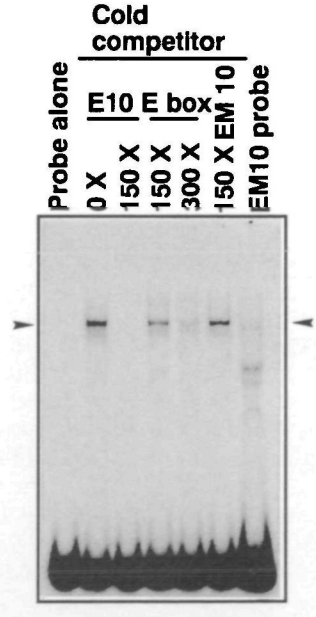

Fig. 7. **Gel mobility shift analysis of nuclear proteins interacting E10 E-box sequences.** Double strand E10 oligonucleotides were used as probe for binding analysis. E10 probe was incubated with nuclear proteins from HepG2 cells in the absence and presence of cold competitors as described under "EXPERIMENTAL PROCEDURES." Competitor oligonucleotides (E10, standard E-box, and EM10) were added at 150–300 molar excess relative to radiolabeled DNA. To confirm specificity of E10 binding protein, radiolabeled EM10 oligonucleotide (EM10 probe) was incubated with nuclear proteins and analyzed.

E-box is a binding motif for various bHLH and bHLH-Zip factors. Such bHLH-Zip factors as Myc (*36*), TFEB (*37*), TFE3 (*38*), and MAD (*39*) have been shown to bind to the E-box motif, CACGTG (*40*). These sequences of E2 (positive element in HepG2 cells), E4-E9 (negative elements in HepG2 and GH4C1 cells), and E10 (positive element in both cells) E-box motifs are CAGGTG, CACCTG and CATTTG, respectively. The sequence CACGTG is not contained in these regulatory E-box elements, suggesting that bHLH-Zip factors are not binding protein for these elements. Recent studies suggest that tissue-specific genes are likely candidates for timely regulation by bHLH transcription factors (*25, 26*). Various bHLH factors such as *ACHAETE SCUTE* genes are involved in neural development, although these targeting genes have not been identified (*41, 42*). Moreover, the bHLH factors have particular importance in the control of peptide hormone synthesis and secretion. Polypeptide hormones including insulin, gastrin, and secretin appear to utilize bHLH factors that are restricted to their differentiated tissues (*43-45*). PACE4 is expressed in some endocrine and neuroendocrine tissues such as the anterior pituitary and central nervous system. Properly functioning bioactive peptide-secreting cells require coordinated expression of hormones or growth factors and their respective processing enzymes. These facts and the present results suggest that bHLH factors are the most likely transcription factors to be involved in the regulation of the PACE4 gene. Our results raise the question of whether bHLH factors involved in regulation of PACE4 gene expression in HepG2 and GH4C1 cells are common or distinct proteins. GH4C1 is a pituitary cell line, and it is known that various bHLH factors are expressed in the pituitary (*46*). However, liver-specific bHLH factors have not been reported yet. BAP (B-activating protein), a rat liver protein that activates transcription through a promoter element (CACGTG) that is similar to the USF/MLTF binding site, was identified by Kugler *et al.* (*47*). However, it is not clear whether BAP is a hepatocyte-specific or a ubiquitous protein. Recently, Aoki *et al.* cloned a novel nuclear protein (RP58), which has a zinc finger DNA binding motif, from spleen cDNA library (*48*). They showed that RP58 binds to the 10 bp-DNA sequence, (A/C)-**ACATCTG**(G/T)(A/C) containing consensus E-box motif and its binding is highly specific. RP58 was shown to repress transcription from a promoter linked to its target sequences. RP58 protein was expressed in the brain at a high level. However, it is unlikely that RP58 is involved in negative regulation of PACE4 gene expression because the sequence of elements for RP58 is quite different from that around E9 E-box, TT**CACCTG**C(A/G). Further analysis is necessary to understand the involvement of RP58 in the regulation of PACE4 gene expression.

Recently, Constam and Robertson (*16*) showed that proBMP is processed to active mature BMP by PACE4 as well as furin and PC6 (*49*). In addition, TGFβ1 (*50*) and activin (*51*) are also shown to be activated by furin. A common feature of all TGFβ growth factors including BMP is that their active forms are dimers of a C-terminal fragment cleaved from a larger precursor (*52*). Generally, the sequence of these processing sites is Arg-X-Arg/Lys-Arg, which is a potential cleavage motif for SPC family proteases. Regulation of BMP activity is probably controlled by members of the SPC family such as PACE4, furin, and PC5/6. BMPs have especially important roles in embryogenesis. Constam *et al.* showed that BMP2, 4, and 7 are coexpressed with PACE4 in the primitive heart, in the apical ectodermal ridge of developing limb buds, and in the interdigital mesenchyme of embryogenic limbs (*12*). A gradient of BMP activity controls the fate of the mediolateral mesoderm cells (*53*). These findings suggest the importance of PACE4 in the regulation of BMP signaling in embryogenesis. On the other hand, several bHLH factors appear to have roles in determining the temporal and spational differentiation of mesoderm. For example, the bHLH protein *Twist* is expressed during the earliest stages of differentiation of the *Drosophila* mesoderm (*54*). It is responsible, in conjunction with the maternal transcription factor *Dorsal*, for activating and maintaining expression of early mesoderm-specific genes including *Snail*. Moreover, bHLH transcription factors play an important role in BMP-4 gene expression (*55*). Similarly, in the case of peptide hormone secretion in endocrine cells, coordinated expression of BMPs and their respective processing enzymes is required for the proper function of BMPs. Several bHLH factors might control BMPs activity by two mechanisms including enhancement of BMP and its processing enzyme expression.

Finally, we showed that bHLH transcription factors have key roles in the regulation of PACE4 expression in HepG2 and GH4C1 cells. Although it remains to be clarified what kinds of transcription factors regulate PACE4 gene expression, our results and recent findings suggest that PACE4 has critical roles in various cell differentiation states such as embryogenesis and neurogenesis, controlled under a dynamic bHLH network.

We wish to thank Professor Shozo Yamamoto and Dr. Kei Yamamoto of the Department of Biochemistry in our University School of Medicine for valuable discussion and technical advice regarding the luciferase assay. We also thank Messrs.Takashi Nishidai and Atsushi Matsuoka for technical assistance, and Ms. Yuka Sasaki for secretarial assistance.

REFERENCES

1. Seidah, N.G., Chretien, M., and Day, R. (1994) The family of subtilisin/kexin like pro-protein and pro-hormone convertases: Divergent or shared functions. *Biochimie* **76**, 197-209
2. Nakayama, K. (1997) Furin: a mammalian subtilisin/Kex2p-like endoprotease involved in processing of a wide variety of precursor proteins. *Biochem. J.* **327**, 625-635
3. Smmekens, S.P. and Steiner, D.F. (1990) Identification of a human insulinoma cDNA encoding a novel mammalian protein structurally related to the yeast dibasic processing protease Kex2. *J. Biol. Chem.* **265**, 2997-3000
4. Smeekens, S.P., Avruch, A.S., LaMendors, J., Chan, S.J., and Steiner, D.F. (1991) Identification of a cDNA encoding a second putative prohormone convertase related to PC2 in AtT20 cells and islets of Langerhans. *Proc. Natl. Acad. Sci. USA* **88**, 340-344
5. Kiefer, M.C., Tucker, J.E., Joh, R., Landsberg, K.E., Saltman, D., and Barr, P.J. (1991) Identification of a second human subtilisin-like protease gene in the fes/fps region of chromosome 15. *DNA Cell Biol.* **10**, 757-769
6. Tsuji, A., Higashine, K., Hine, C., Mori, K., Tamai, Y., Nagamune, H., and Matsuda, Y. (1994) Identification of novel cDNAs encoding human kexin-like protease, PACE4 isoforms. *Biochem. Biophys. Res. Commun.* **200**, 945-950
7. Mori, K., Kii, S., Tsuji, A., Nagahama, M., Imamaki, A., Hayashi, K., Akamatsu, T., Nagamune, H., and Matsuda, Y. (1997) A novel human PACE4 isoform, PACE4E is an active

processing protease containing a hydrophobic cluster at the carboxy terminus. *J. Biochem.* **121**, 941-948
8. Nakayama, K., Kim, W.-S., Torii, S., Hosaka, M., Nakagawa, T., Ikemizu, J., Baba, T., and Murakami, K. (1992) Identification of the fourth member of the mammalian endoprotease family homologous to the yeast Kex2 protease: Its testis-specific expression. *J. Biol. Chem.* **267**, 5897-5900
9. Nakagawa, T., Hosaka, M., Torii, S., Watanabe, T., Murakami, K., and Nakayama, K. (1993) Identification and functional expression of a new member of the mammalian Kex2-like processing endoprotease family: Its striking structural similarity to PACE4. *J. Biochem.* **113**, 132-135
10. Bruzzaniti, A., Goodge, K., Jay, P., Taviaux, S.A., Lam, M.H.C., Berta, P., Martin, T.J., Moseley, J.M., and Gillespie, M.T. (1996) C8, a new member of the convertase family. *Biochem. J.* **314**, 727-731
11. Meerabux, J., Yaspo, M.L., Roebroek, A.J., Van de Ven, W.J., Lister, T.A., and Yaung, B.D. (1996) A new member of the proprotein convertase gene family (LPC) is located at a chromosome translocation breakpoint in lymphomas. *Cancer Res.* **56**, 448-451
12. Constam, D.B., Calfon, M., and Robertson, E.J. (1996) SPC4, SPC6, and the novel protease SPC7 are coexpressed with bone morphogenetic proteins at distinct sites during embryogenesis. *J. Cell. Biol.* **134**, 181-191
13. Nagamune, H., Muramatsu, K., Akamatsu, T., Tamai, Y., Izumi, K., Tsuji, A., and Matsuda, Y. (1997) Distribution of the kexin family proteases in pancreatic islets: PACE4C is specifically expressed in B cells of pancreatic islets. *Endocrinology* **136**, 357-360
14. Dong, W., Marcinkiewicz, M., Vieau, D., Chrétien, M., Seidah, N.G., and Day, R. (1995) Distinct mRNA expression of the highly homologous convertases PC5 and PACE4 in the rat brain and pituitary. *Neuroscience* **15**, 1778-1796
15. Akamatsu, T., Daikoku, S., Nagamune, H., Yoshida, S., Mori, K., Tsuji, A., and Matsuda, Y. (1997) Developmental expression of a novel Kexin family protease, PACE4E, in the rat olfactory system. *Histochem. Cell. Biol.* **108**, 95-103
16. Constam, D.B. and Robertson, E.J. (1999) Regulation of bone morphogenetic protein activity by pro domains and proprotein convertases. *J. Cell. Biol.* **144**, 139-149
17. Mori, K., Imamaki, A., Nagata, K., Yonetomi, K., Kiyokage-Yoshimoto, R., Martin, T.J., Gillespie, M.T., Nagahama, M., Tsuji, A., and Matsuda, Y. (1999) Subtilisin-like proprotein convertases, PACE4 and PC8, as well as furin, are endogenous proalbumin convertases in HepG2 cells. *J. Biochem.* **125**, 627-633
18. Tsuji, A., Hine, C., Tamai, Y., Yonemoti, K., Mori, K., Yoshida, S., Bando, M., Sakai, E., Mori, K., Akamatsu, T., and Matsuda, Y. (1997) Genomic organization and alternative splicing of human PACE4 (SPC4), kexin-like processing endoprotease. *J. Biochem.* **122**, 438-452
19. Roebroek, A.J.M., Schalken, J.A., Leunissen, J.A.M., Onnekink, C., Bloemers, H.P.J., and Van de Ven, W.J.M. (1986) Evolutionary conserved close linkage of the c-fes/fps proto-oncogene and genetic sequences encoding a receptor-like protein. *EMBO J.* **5**, 2197-2202
20. Barr, P.J., Mason, O.B., Landsberg, K.E., Wang, P.A., Keifer, M.C., and Brake, A.J. (1991) cDNA and gene structure for a human subtilisin-like protease with cleavage specificity for paired basic amino acid residues. *DNA Cell Biol.* **10**, 319-328
21. Ohagi, S., LaMendola, J., LeBeau, M.M., Epinosa III, R., Takada, J., Smeekens, S.P., Chan, S.J., and Steiner, D.F. (1992) Identification and analysis of the gene encoding human PC2, a prohormone convertase expressed in neuroendocrine tissues. *Proc. Natl. Acad. Sci. USA* **89**, 4977-4981
22. Ftouhi, N., Day, R., Mbikay, M., Crétien, M., and Seidah, N.G. (1994) Gene organization of the mouse pro-hormone and pro-protein convertase PC1. *DNA Cell Biol.* **13**, 395-407
23. Mbikay, M., Raffin-Saison, M.-L., Tadros, H., Sirois, F., Seidah, N.G., and Crétien, M. (1994) Structure of the gene for the testis-specific proprotein convertase 4 and of its alternate messenger RNA isoforms. *Genomics* **20**, 231-237
24. Goodge, K.A., Thomas, R.J., Martin, T.J., and Gillespie, M.T. (1998) Gene organization and alternative splicing of human prohormone convertase PC8. *Biochem. J.* **336**, 353-359
25. Jan, Y.N. and Jan, L.Y. (1993) HLH proteins, fly neurogenesis and vertebrate myogenesis. *Cell* **75**, 827-829
26. Campos-Ortega, J.A. and Jan, Y.N. (1991) Genetic and molecular bases of neurogenesis in Drosophila melanogaster. *Annu. Rev. Neurosci.* **14**, 399-420
27. Chiaramello, A., Neuman, T., Deavy, D.R., and Zuber, M.X. (1996) The GAP-43 gene is a direct downstream target of the basic helix-loop-helix transcription factors. *J. Biol. Chem.* **271**, 22035-22043
28. Higuchi, R., Krummel, B., and Saiki, R.K. (1988) A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Res.* **16**, 7351-7367
29. Arakawa, T., Nakamura, M., Yoshimoto, T., and Yamamoto, S. (1995) The transcriptional regulation of human arachidonate 12-lipoxygenase gene by NF kappa B/Rel. *FEBS Lett.* **363**, 105-110
30. Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**, 1475-1489
31. Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248-254
32. Miyagoe, Y., Galibert, M.-D., Georgatsou, E., Fourel, G., and Meo, T. (1994) Promoter elements of the mouse complement C4 gene critical for transcription activation and start site location. *J. Biol. Chem.* **269**, 8268-8279
33. Pelletier, M.R., Hatada, E.N., Scholz, G., and Scheidereit, C. (1997) Efficient transcription of an immunoglobulin κ promoter requires specific sequence elements overlapping with and downstream of the transcriptional start site. *Nucleic Acids Res.* **25**, 3995-4003
34. Jansen, E., Ayoubi, T.A., Meulemans, S.M.P., and Van de Ven, W.J.M. (1995) Neuroendocrine-specific expression of the human prohormone convertase 1 gene. Hormonal regulation of transcription through distinct cAMP response elements. *J. Biol. Chem.* **270**, 15391-15397
35. Jansen, E., Ayoubi, T.A., Meulemans, S.M.P., and Van de Ven, W.J.M. (1997) Cell type-specific protein-DNA interactions at the cAMP response elements of the prohormone convertase 1 promoter. Evidence for additional transactivators distinct from CREB/ATF family members. *J. Biol. Chem.* **272**, 2500-2508
36. Torres, R., Schreiber-Agus, N., Morgenbesser, S.D., and DeOinho, R.A. (1992) Myc and Max: a putative transcriptional complex in search of a cellular target. *Curr. Opin. Cell. Biol.* **4**, 468-474
37. Carr, C.S. and Sharp, P.A. (1990) A helix-loop-helix protein related to the immunoglobulin E box-binding proteins. *Mol. Cell. Biol.* **10**, 4384-4388
38. Beckman, H., Su, L.-K., and Kadesch, T. (1990) TFE3: a helix-loop-helix protein that activates transcription through the immunoglobulin enhancer muE3 motif. *Genes Dev.* **4**, 167-179
39. Ayer, D.E., Kretzner, L., and Einenman, R.N. (1993) Mad: a heterodimeric partner for Max that antagonizes Myc transcriptional activity. *Cell* **72**, 211-222
40. Bendall, A.J. and Molloy, P.L. (1994) Base preferences for DNA binding by the bHLH-Zip protein USF: effects of $MgCl_2$ on specificity and comparison with binding of Myc family members. *Nucleic Acid Res.* **22**, 2801-2810
41. Campuzano, S. and Modolell, J. (1992) Patterning of the *Drosophila* nervous system: the achaete-scute gene complex. *Trends Genet.* **8**, 202-208
42. Guillemot, F., Lo, L.-C., Johnson, J.E., Auerbach, A., Anderson, D.J., and Joyner, A.L. (1993) Mammalian achaete-scute homolog 1 is required for the early development of olfactory and autonomic neurons. *Cell* **75**, 463-476
43. Moss, L.G., Moss, J.B., and Rutter, W.J. (1988) Systematic

binding analysis of the insulin gene transcription control region: insulin and immunoglobulin enhancers utilize similar transactivators. *Mol. Cell. Biol.* **8**, 2620-2627
44. Wang, T.C. and Brand, S.J. (1990) Islet cell-specific regulatory domain in the gastrin promoter contains adjacent positive and negative DNA elements. *J. Biol. Chem.* **265**, 8908-8914
45. Wheeler, M.B., Nishitani, J., Buchnan, A.M.J., Kopin, A.S., Chey, W.Y., Chang, T.-A., and Leiter, A.B. (1992) Identification of a transcriptional enhancer important for enteroendocrine and pancreatic islet cell-specific expression of the secretin gene. *Mol. Cell. Biol.* **12**, 3531-3539
46. Jackson, S.M., Barnhart, K.M., Mellon, P.L., Gutierrez-Hartmann, A., and Hoeffler, J.P. (1993) Helix-loop-helix proteins are present and differentially expressed in different cell lines from the anterior pituitary. *Mol. Cell. Endocrinol.* **96**, 167-176
47. Kugler, W., Kaling, M., Ross, K., Wagner, V., and Ryffel, G.U. (1990) BAP, a rat liver protein that activates transcription through a promoter element with similarity to the USF/MLTF binding site. *Nucleic Acids Res.* **18**, 6943-6969
48. Aoki, K., Meng, G., Suzuki, K., Takashi, T., Kameoka, Y., Nakahara, K., Ishida, R., and Kasai, M. (1998) RP58 associates with condensed chromatin and mediates a sequence-specific transcriptional repression. *J. Biol. Chem.* **273**, 26698-26704
49. Cui, F., Jean, F., Thomas, G., and Christian, J.L. (1998) BMP-4 is proteolytically activated by furin and/or PC6 during vertebrate embryonic development. *EMBO J.* **17**, 4735-4743
50. Dubois, C.M., Laprise, M.H., Blanchette, F., Gentry, L.E., and Leduc, R. (1995) Processing of transforming growth factor beta 1 precursor by human furin convertase. *J. Biol. Chem.* **270**, 10618-10624
51. Roebroek, A.J.M., Creemers, J.W.M., Pauli, I.G.L., Bogaert, T., and Van de Ven, W.J.M. (1993) Generation of structural and functional diversity in furin-like proteins in *Drosophila melanogaster* by alternative splicing of the Dfur1 gene. *EMBO J.* **12**, 1853-1870
52. Massague, J. (1987) The TGF-beta family of growth and differentiation factors. *Cell* **49**, 437-438
53. Tonegawa, A., Funayama, N., Ueno, N., and Takahashi, Y. (1997) Mesodermal subdivision along the mediolateral axis in chicken controlled by different concentrations of BMP-4. *Development* **124**, 1975-1984
54. Reuter, R. and Leptin, M. (1994) Interacting functions of snail, twist and huckebein during the early development of germ layers in Drosophila. *Development* **120**, 1137-1150
55. Ebara, S., Kawasaki, S., Nakamura, I., Tsutsumimoto, T., Nakayama, K., Nikaido, T., and Takaoka, K. (1997) Transcriptional regulation of the mBMP-4 gene through an E-box in the 5′-flanking promoter region involving USF. *Biochem. Biophys. Res. Commun.* **240**, 136-141