Constitutive Expression of the Gene for the Cell-specific p48 DNA-binding Subunit of Pancreas Transcription Factor 1 in Cultured Cells Is under Control of Binding Sites for Transcription Factors Sp1 and αCbf*

(Received for publication, January 24, 1996, and in revised form, April 19, 1996)

Martin Knöffler†, Andrea Krapp, Otto Hagenbüchle, and Peter K. Wellauer‡
From the Swiss Institute for Experimental Cancer Research (ISREC), CH-1066 Epalinges/Lausanne, Switzerland

We have cloned and characterized the rat gene that encodes the p48 DNA-binding subunit of pancreas transcription factor 1 (Ptf1), a cell-specific basic region helix-loop-helix (bHLH) protein. The ptf1-p48 gene measures 1.8 kilobases in size and occurs as a single copy in the haploid genome. Run-on transcription assays suggest that this gene is subject to transcriptional control since no activity of its promoter is detected in non-producing cells. The gene specifies two mRNAs that encode the same protein and originate from transcription initiation at alternative sites. Expression analysis of hybrid genes bearing deletions of the gene’s 5’-flanking region fused to a reporter gene defines a promoter region within the gene-proximal 260 base pairs of DNA. The cis-acting elements that control promoter activity include binding sites for transcription factors Sp1 and αCbf, a 60-kDa CCAAT box-binding protein. The gene promoter, however, functions not only in exocrine pancreatic cells but also in cells of other origin. No cell-specific transcriptional control element was detected in as much as 10 kilobases of 5’-flanking region. We discuss models of how the cell-specific expression of the endogenous ptf1-p48 gene might be established during development of the animal.

The expression of genes coding for the cell-specific products of terminally differentiated cells is often under the control of cell-specific transcription factors. However, the mechanisms which determine that cell-specific regulatory proteins are synthesized in a correct spatial and temporal order during the development of a multicellular organism are still poorly understood. One approach to study this problem is to explore the regulatory circuits that govern the expression of genes encoding such factors. An important question that may be addressed by a developmental “regression” analysis is whether regulatory genes that are expressed in a cell-specific fashion are themselves subject to regulation by cell lineage-specific transcription factors. There is so far no compelling evidence either for or against a general role of cascades composed of cell lineage-specific transcription factors. One reason for this may be that the key regulators which transactivate the genes for cell-specific transcription factors have so far proven quite elusive. For instance, the cis-acting DNA elements and protein factors required for correct temporal transactivation of the genes encoding muscle-specific transcription factors (1–4) during development are still largely unknown despite considerable efforts to identify them. It is therefore believed that the factors required for activation of gene expression during development are not necessarily those which ensure maintenance of transcription later on (5–7).

We have studied the expression of a gene that encodes a cell-specific DNA-binding subunit of Ptf1, an enhancer-binding protein that coordinately regulates the expression of genes encoding the specific functions of the exocrine pancreas (8). Ptf1 binding activity is first detected at day 15 of mouse development concomitant with the start of efficient transcription of its target genes (9, 10). Ptf1 is a heterooligomer containing three distinct protein subunits. Two of these, p48 and p64, contact the DNA, while the third one, p75, does not but is required for import of the factor into the cell nucleus (11). The p48 and p64 subunits do not bind individually but contact a bipartite binding site that encompasses two distinct recognition sequences (8, 12). The p48 subunit is a cell-specific member of the bHLH* class of proteins. Its presence is restricted to cells of the exocrine pancreas both in the adult animal and the embryo (9). Expression of p48 antisense RNA in exocrine pancreatic cells in culture down-regulates transcriptional activity of Ptf1-dependent genes suggesting that the protein is critically involved in the maintenance of exocrine pancreas-specific gene expression. However, unlike myogenic factors, p48 is incapable of establishing a cell-specific transcription program on its own when introduced into cell lines of nonpancreatic origin.

Here we report the cloning and characterization of the gene encoding the p48 subunit of Ptf1 and study its expression. We show that this gene contains, within the gene-proximal 260 bp of 5’-flanking region, a promoter that is active both in expressing and nonexpressing cells. The cis-acting elements responsible for this promoter activity include multiple binding sites for transcription factor Sp1 and the recognition sequence for the CCAAT box binding factor αCbf.

EXPERIMENTAL PROCEDURES

Screening of a Genomic DNA Library and Analysis of λ Phage Inserts—A λ DASH II genomic DNA library of Sprague-Dawley rats (Stratagene) was screened with full-length p48 cDNA (9) by in situ plaque hybridization. Briefly, 10⁶ λ plaques were transferred to nitro...
Transcription Control of the ptf1-p48 Subunit Gene

cellulose membranes (Schleicher and Schuell) and hybridized with 32P-labeled cDNA in a buffer containing 6 x SSC, 0.1% Nonidet P-40, and 0.05 x Biotto (13). A total of 11 positive clones was obtained. λ-DNA from individual clones was purified by the plate lysate method and digested with EcoRI. Restriction fragments were separated by agarose gel electrophoresis and transferred to nylon membranes (GeneScreen, Dupont NEN) using standard methods. All 11 clones were found to be independent and carried inserts ranging from about 11 to 20 kb in size. Hybridization with 32P-labeled DNA occurred to a single EcoRI fragment which, in 6 out of the 11 clones, measured 11.5 kb in size. This fragment was cloned in hybridization with strand-separating genomic DNA fragments. Hybridization was found to be exclusive to the coding DNA strand. The origin of RNA probes as polymerase II transcripts was established by in vitro elongation reactions carried out in the presence of 2 μg/ml a-amanitin.

Construction of Hybrid Genes—A 4.5-kb restriction fragment of recombiant phage ptf1-p48 carrying a unique SalI site (position 1, and the left end of the SalI fragment of ptf1-p48 genomic DNA was subcloned into the XbaI site of pGEM-3. The fragment was recovered from the recombinant plasmid by digestion with XbaI, gel-purified and cleaved at a Stul site within exon 1. The resulting 1.1-kb fragment (+216 to −900) was ligated into the Smal site of the luciferase reporter vector pGL2 (Promega) to yield hybrid gene B of Fig. 4a. Hybrid gene C was made by adding SalI linkers to a gel-purified 670-bp Smal fragment of subcloned genomic DNA (+86 to −591) which was then ligated into the Xhol site of pGL2. Hybrid genes A, D, and E of Fig. 4a were constructed from hybrid gene C by digestion with appropriate combinations of restriction endonucleases. Hybrid gene D was made by partially cleaving DNA of C with EcoRI and then with Smal at the unique site of the pGL2 palinker. The DNA fragment having the correct length for one cleaved with EcoRI at position −258 was gel-purified and self-ligated to make hybrid gene E, DNA of C was digested with BssHII (at position −64) and Smal (within the palinker). The vector containing DNA fragment was blunt-ended, gel-purified, and self-ligated. Hybrid gene A containing ptf1-p48 gene sequences +86 to −3600 was obtained by digesting DNA of C with XbaI (at position −64) and then using the XbaI site of pGL2 palinker. A 3.6-kb SacI/BssHII fragment of a recombinant plasmid carrying 9.7 kb of genomic DNA (nucleotides +86 to −960) was then inserted into the vector by forced cloning. Hybrid genes I-1 and I-2 were constructed from hybrid gene C digested with appropriate combinations of restriction endonucleases and labeled with 32P end-labeled T4 DNA polymerase. The DNA fragment was then ligated with NaeI and then cleaved with EcoRI and MspI.

In vitro Mutagenesis Reactions—DNA fragments containing point mutations were excised from the vector with SalI and ligated in their sense orientation into the Xhol site of pGEM-3. For stable transfection of cells in culture, plasmid DNAs were linearized at the SalI site of pGEM-3.

Tissue Culture—Pancreatic rat AR42J and AR4IP cells were cultured at 37°C and 5% CO2 in a medium containing 7% fetal calf serum. Mouse NIH3T3 cells were made by Life Technologies, Inc.), 25% F12, 25% Dulbecco’s modified Eagle’s medium and 10% fetal calf serum. Mouse NIH3T3 cells were harvested by trypsinization, washed twice with phosphate-buffered saline, and resuspended in culture medium at a concentration of 105 cells/800 μl. 10 μg of linearized hybrid gene DNA and 1 μg of linearized pHMG1 (Stratagene) were added to the cell suspension on ice. Transfections of AR42J and NIH3T3 cells were performed by electroporation at 900 microfarads in a Gene Pulser apparatus (Bio-Rad) at 350 V and 280 V, respectively. AR42J and NIH3T3 cells were subjected to neomycin selection for 10 days in culture media containing the drug G418 (Geneticin; Life Technologies, Inc.) at a concentration of 260 μg/ml and 500 μg/ml, respectively. Pools of neomycin-resistant cell clones were grown to confluency on duplicate Petri dishes for DNA isolation and enzymatic assays, respectively. Protein extracts and luciferase assays were made using a commercial luciferase assay system (Promega) under conditions specified by the supplier. Luciferase activity was quantified in a Biocounter 2500 luminometer. DNA extraction was performed by SDS-lysis of protein with phenol-chloroform. DNA hybridization signals were quantified by analysis in a PhosphorImager (Molecular Dynamics). Luciferase activity was standardized by measuring protein with a Bio-Rad protein assay reagent (Bio-Rad) and then normalized for the amount of integrated luciferase DNA.

DNase I footprinting. Electrophoretic Mobility Shift Assay (EMSA), and UV-crosslinking—For footprint analysis, a plasmid containing ptf1-p48 sequences +86 to −591 was cleaved at the EcoRI site (position −258), and the coding strand was 32P-end-labeled using T4 DNA polymerase. The DNA was then cleaved with NaeI (position +86), and the 250-bp fragment was gel-purified. For labeling of the noncoding strand, the plasmid was first digested at the Nael site, labeled, and then cleaved with EcoRI. Nuclear protein extracts (N.E.) were prepared from AR42J and AR4IP cells by a modification of the procedure of Ref. 20 as described earlier (8). Binding reactions with nuclear extract, DNAase I digestion, and gel electrophoresis were done according to Ref. 8. EMSA for the detection of protein-DNA complexes was carried out by electrophoresis of binding reactions on 2% agarose in 0.5 x Tris borate EDTA buffer (pH 8.3) at 40 mA in the cold. Binding reactions (20 μl) were done for 5 min at room temperature in a mixture containing 0.1 pmol of double-stranded, 32P-labeled oligonucleotide, 7 ng of N.E., and 1 μg of poly(dI·dC). Binding for the detection of Sp1 in N.E. was carried out in 15% glycerol, 65 mM KCl, 11 mM HEPES (pH 7.9), 2 mM dithiorthreitol, 0.5 mM EDTA, and 0.1 mM ZnCl2 (buffer A). Binding reactions for detection of C1 and 2, C/EBP, and αCf in N.E. were carried out in buffer A lacking ZnCl2, and, as a control, in buffer A containing 5 mM MgCl2. No qualitative, but some minor quantitative differences in binding were observed under these experimental conditions. The 32P-labeled purified human Sp1 (Promega) was done in a buffer containing 12% glycerol, 55 mM KCl, 12 mM HEPES (pH 7.9), 4 mM dithiorthreitol, 0.5 mM EDTA, 0.2 mM ZnCl2, 2 mM MgCl2, and 350 μg of bovine serum albumin/ml. Radiolabeled protein-DNA complexes in various binding reactions were quantified in a PhosphorImager. The sequence specificity of all protein-DNA complexes reported in this paper was determined by competition with heterologous DNA sequences.

UV-crosslinking of protein-DNA complexes was done essentially as described in Ref. 12. The radiolabeled, double-stranded oligonucleotide bound to Sp1, in which T residues were replaced by azidodeoxuryridine (N3dU), was synthesized by annealing the coding DNA strand bearing the CCAAT sequence with a short noncoding strand primer that was extended with Klenow DNA polymerase in the presence of N3dUTP and N3dC (1000 Ci/mmol). Binding reactions were carried out in buffer A lacking ZnCl2, but containing 1 mg/ml single-stranded Escherichia coli DNA as a competitor for nonspecific DNA-binding proteins. Cross-linked protein-DNA complexes were separated on 10% SDS-polyacrylamide gels.

RESULTS

We have isolated the gene encoding the mRNA for an exocrine pancreas-specific bHLH protein, the p48 DNA-binding subunit of transcription factor Ptf1. Screening of a rat genomic λ library with full-length p48 DNA (9) yielded several positive clones which, when analyzed by restriction enzyme digestion, were found to share common DNA fragments. One of these clones bearing a copy of the gene on a 11.5-kb EcoRI fragment was chosen for further analysis and subjected to heteroduplex analysis in the electron microscope (Fig. 1, a and b) or restric-
tion enzyme mapping (Fig. 1c). Heteroduplexes between p48 cDNA and the 11.5-kb EcoRI fragment show that the genomic DNA contains both gene and flanking region sequences. The heteroduplex delimits a gene region of about 1.8 kb in size containing two exons and a small intron (Fig. 1, a and b). The evidence from Southern hybridization of genomic DNA (Fig. 1d) is compatible with the existence of a single copy of this gene per haploid rat genome since DNA fragments hybridizing to the p48 cDNA probe are those predicted from the restriction map of the genomic clone. The data do not formally exclude the remote possibility, however, that additional copies of the gene, producing the same restriction pattern over a large stretch of DNA, may occur elsewhere in the genome. DNA sequence analysis of the gene (Fig. 2) confirms the general architecture deduced from electron microscopy. The single intron, 331 bp in size, is located at nucleotide position 11015 of the gene sequence. Exon 1 thus determines 5'-nontranslated and N-terminal protein coding sequences including the bHLH domain, while exon 2 specifies the rest of the protein coding region and 3'-translated sequences of p48 mRNA. The ptf1-p48 gene produces two mRNA species that are 1.5 kb and 1.3 kb in size and differ in length of their 5'-translated regions (9). Comparison of genomic DNA and p48 cDNA sequences (9) shows that they are colinear. This is consistent with the idea that the two p48 mRNA species originate from the same gene by transcription initiation at two alternative sites. A TATA-element (position 227) precedes the cap site (11L) for the 1.5-kb mRNA, and a related CACA motif is located 28 bp upstream of the cap site (11S) for the 1.3-kb mRNA species. The evidence from run-on transcription assays shows that...
transcription initiation in vivo occurs within a region of genomic DNA that specifies the two cap sites (fragment b in Fig. 3a). The combined transcriptional activity originating from the two cap sites is specific for cells synthesizing exocrine pancreatic products, such as the rat AR42J cell line (Fig. 3b). No transcription is detected in nuclei of AR4IP cells that derive from the same transplantable tumor as AR42J cells (21) but do not express differentiated functions (Fig. 3b). These observations indicate that transcription of the ptf1-p48 gene is subject to cell-specific control.

To determine the region of genomic DNA that governs ptf1-p48 gene expression, we have constructed a series of hybrid genes containing a luciferase reporter gene under control of different portions of the ptf1-p48 5'-flanking region (Fig. 4a). The relative transcriptional activity of various hybrid genes was determined indirectly by measuring enzymatic activity produced by the reporter gene in stably transfected AR42J and mouse NIH3T3 cells (NIH3T3 cells, like AR4IP cells, do not express the endogenous ptf1-p48 gene but prove far superior for transfection). Luciferase activity was determined in pools containing a large number of individual cellular clones since, on the average, only one copy of the transgene integrated per cell (data not shown). The results of two independent transfection experiments for each cell line are compiled in Fig. 4b and show that gene-proximal 5'-flanking sequences present in hybrid gene D illicit a 36- and 52-fold stimulation of reporter gene activity over basal levels in AR42J and NIH3T3 cells, respectively (Fig. 4a). The DNA sequences between −65 and −258 thus contain cis-acting DNA elements that increase the efficiency of the residual (TATA box) promoter (+1 to −64). Neither these, nor DNA sequences in the remaining part of the 3.6-kb 5'-flanking region confer cell-specific expression to the reporter gene, however, since all constructs are active in both cell types. For reasons not understood, DNA sequences located between −900 and −3600 reproducibly repressed reporter gene expression in the two cell lines (compare activities of hybrid gene A with those of C or D in Fig. 4a). Insertion of additional 5.8 kb of 5'-flanking sequences (−3600 to −9400) into hybrid gene A did not relieve this repression in either cell type (data not shown). A 2- to 3-fold higher expression occurs with hybrid gene B as compared to hybrid genes C or D (Fig. 4a). This may be explained by the existence of additional control elements that positively affect transcription within the region comprised between −591 and −900 and/or the presence of the cap site (+15) for the smaller p48 transcript. The fact that the two mRNA species produced by the ptf1-p48 gene reside in about equimolar amounts inside the exocrine pancreatic cell (9) favors the second alternative.

In order to identify trans-acting factors responsible for the expression of hybrid genes in cultured cells, we have searched for sites of protein-DNA interaction within the gene-proximal 260 bp of the 5'-flanking region by carrying out in vitro DNase I footprinting with nuclear proteins of AR42J or AR4IP cells (Fig. 5). Five distinct domains of an end-labeled DNA fragment bearing sequences −8 to −258 were found to be protected in a reproducible manner on both strands. There is no evidence for cell-specific sites of protein-DNA interactions since a common pattern of protection is observed irrespective of the origin of nuclear protein used. Several of the protected DNA domains (designated I to IV and TATA in Fig. 5) encompass DNA sequences that constitute putative binding sites for known transcription factors. Protection of the TATA region is expected to result from binding of TFIIID complex (22). DNA sequences specifying domains I, II, III, and IV all encompass elements sharing similarities with binding sites for transcription factor Sp1 (23, 24). The DNA of footprint domain I includes a CCAAT box at position −68 suggesting that protection in this case results from interaction with a member of the family of proteins recognizing this particular sequence motif.

We have tested the assumption that the various DNA sequences protected by nuclear protein constitute binding sites for sequence-specific DNA-binding proteins by carrying out electrophoretic mobility shift and competition assays. Synthetic oligonucleotides harboring sequences of footprints II, III, and IV (Fig. 6a) were incubated with N.E. of AR42J cells to explore their potential for the binding of transcription factor Sp1. The results from these binding reactions indicate that all three sequences recognize, albeit with different relative affinities, Sp1 binding activity. The three protein-DNA complexes generated are sequence-specific and indistinguishable from those formed with an oligonucleotide bearing a canonical Sp1 recognition motif, and their formation is efficiently inhibited by a large excess of cold Sp1 sequence (Fig. 6b). However, higher molar ratios of II, III, and IV sequence are required for competition of these complexes as compared to homologous sequence suggesting that the various DNAs bind Sp1 activity(ies) of AR42J cells with differential affinities (Fig. 6c). Quantification of protein-DNA complexes by phosphorimage analysis (not shown) establishes a hierarchy of binding affinities of these
DNAs with Sp1 > II > IV > III. The same order of relative affinities is also detected in binding reactions with purified human SP1 protein. This protein generates a single complex with SP1, II, or IV DNA, interacts inefficiently with III DNA, and, as expected, fails to recognize heterologous I sequence (Fig. 6d). The TATA box upstream of +1 is at nucleotide position −27. The numbers on the right represent the activity of hybrid genes A–D relative to that of hybrid gene E (arbitrarily taken as 1) as calculated from the data shown in b. The expression of hybrid genes A, C, D, and E in AR42J and NIH3T3 cells was determined by measuring the amount of luciferase activity in two independent transfection experiments (I and II) for each cell line. The activity of hybrid gene B was determined in a single experiment for each cell type only. The values determined for all hybrid genes are shown in the bar diagram of b and are based on measurements of pools containing a large number (40–150) of independent AR42J or NIH3T3 clones. The expression of hybrid genes A–E in experiments I and II can be compared directly since enzymatic activity has been normalized for the amount of full-length luciferase DNA present in the cells. The normalized luciferase activity of hybrid gene E was arbitrarily taken as 1. Note, however, that enzymatic activity in AR42J and NIH3T3 cells cannot be compared directly since luciferase mRNA and/or protein may have different stability in the two cell lines.

DNA binding is also dependent on the extent of sequence similarity between the oligonucleotide and the DNA sequence. For example, the TATA box upstream of +1 is at nucleotide position −27. The numbers on the right represent the activity of hybrid genes A–D relative to that of hybrid gene E (arbitrarily taken as 1) as calculated from the data shown in b. The expression of hybrid genes A, C, D, and E in AR42J and NIH3T3 cells was determined by measuring the amount of luciferase activity in two independent transfection experiments (I and II) for each cell line. The activity of hybrid gene B was determined in a single experiment for each cell type only. The values determined for all hybrid genes are shown in the bar diagram of b and are based on measurements of pools containing a large number (40–150) of independent AR42J or NIH3T3 clones. The expression of hybrid genes A–E in experiments I and II can be compared directly since enzymatic activity has been normalized for the amount of full-length luciferase DNA present in the cells. The normalized luciferase activity of hybrid gene E was arbitrarily taken as 1. Note, however, that enzymatic activity in AR42J and NIH3T3 cells cannot be compared directly since luciferase mRNA and/or protein may have different stability in the two cell lines.

Potential candidates for proteins interacting with DNA sequences delimited by footprint I are those that recognize a CCAAT motif, such as NF1, CP1 and -2, C/EBP, or αCBF. Two lines of experimental evidence argue against NF1 (27) as the binding activity of interest. An oligonucleotide bearing a canonical NF1 recognition motif (28) failed to compete with DNA sequences defining footprint I, and anti-NF1 antibody did not recognize complexes containing oligonucleotide I even though it supershifted a complex containing genuine NF1 (data not shown). C/EBP is excluded as the candidate activity since its binding site is unable to compete, even at high molar excess, with two sequence-specific complexes formed with oligonucleotide I (Fig. 7a). CP1 and CP2 are also unlikely candidates since their binding sites compete only at high molar excess and this despite a considerable degree of homology between, for example, CP1 DNA and oligonucleotide I (Fig. 7a). All evidence favors αCBF as the key factor. Not only does its cognate sequence efficiently compete for binding of activity(ies) recognizing oligonucleotide I and vice versa (Fig. 7, a and b), but also the sequence requirements for αCBF binding are in-
FIG. 5. DNase I footprinting of the ptf1-p48 gene promoter. a, a NaeI/Eco47III restriction fragment (nucleotides −8 to −258) in which either the noncoding (nc) or the coding (c) DNA strand had been end-labeled was incubated with N.E. of AR42J or AR41P cells. Protein-DNA complexes were partially digested with DNase I, and digestion products were separated by electrophoresis on a 6% sequencing gel together with those obtained by DNase I treatment of the the DNA alone. G + A cleavage reactions (44) were included as sequence markers. DNA sequences that were protected by nuclear protein on both strands are indicated by brackets on the right side of each autoradiograph and have been termed TATA,
distinguishable from those of protein(s) binding to oligonucleotide I (Fig. 7c). αCBF as well as oligonucleotide I binding activity contact, in addition to the CCAAT box, nucleotides flanking this motif on either side. For instance, point mutants I-PM2 and I-PM3 are 7 and 22 times less efficient in binding, respectively, than wild type sequence. Deletion of a single G residue in I-DM2 reduces binding by a factor of 7, and the even shorter sequence of I-DM1 is deficient for binding altogether (Fig. 7c). In contrast, only minor effects upon binding were observed by replacing nucleotides within the oligonucleotide I sequence that spans the ’3’ end of the CCAAT box and the terminal two G residues (data not shown). This observation suggests that a critical length of the DNA is more important than the actual sequence for factor binding and implies that contacts to the phosphate backbone play a predominant role over those occurring to specific bases in this region of the DNA. One of the proteins that constitute oligonucleotide I binding activity was identified by UV-crosslinking (Fig. 7d). A predominant sequence-specific complex having an apparent molecular mass of about 68 kDa is generated when nuclear protein of AR42J or AR41P cells is cross-linked to radiolabeled N3'dU-substituted oligonucleotide I. Formation of this complex is competed in a reciprocal manner by an excess of cold, nonsubstituted αCBF and I sequence. The complex is expected to harbor a single protein due to the low efficiency of UV-crosslinking (12). If we subtract the molecular mass of cross-linked DNA singlestrand, we estimate the protein to measure about 60 kDa in size. The observation that this protein occurs in expressing as well as nonexpressing cells (Fig. 7d) indicates that it is not cell-specific.

To determine whether the binding sites for Sp1 and αCbf established by protein-DNA binding assays in vitro are functionally relevant, we have made a series of hybrid genes bearing point mutations in individual binding sites (Fig. 8a). These mutants were then stably transfected into AR42J cells. Expression analysis shows that vectors carrying mutations within Sp1 binding sites II and IV or αCbf binding site I produce lower levels of enzymatic activity than wild type sequence (Fig. 8b).

I, II, III, and IV, respectively. Their assignment to the DNA sequence is shown in b. Numbers are nucleotide positions relative to the transcription initiation site for 1.5-kb mRNA (+1L). The TATA element, a CCAAT motif, and putative Sp1 binding sites are boxed (see “Results”). The end at which the DNA was labeled is indicated by an asterisk.

Transcription Control of the ptf1-p48 Subunit Gene

Fig. 6. Identification of binding sites for transcription factor Sp1 in the upstream region of the ptf1-p48 gene. The origin of nuclear proteins of AR42J cells interacting with synthetic oligonucleotides that define footprint domains II, III, and IV was studied by EMSA. Binding reactions were analyzed by electrophoresis on 2% agarose gels. a, the nucleotide sequences of oligonucleotides II, III, and IV and one containing a canonical Sp1 binding site (45) are compared. Putative Sp1 binding sites in oligonucleotides II, III, and IV and the canonical site present in oligonucleotide SP1 are boxed. Note that oligonucleotide IV contains two partially overlapping candidate Sp1 sites (IV-1, IV-2). The homology comparison of putative Sp1 sites with the Sp1 consensus sequence is shown below the oligonucleotide sequences. Nucleotides of II, III, and IV that fit the consensus are boxed. Point mutations (PM) were introduced into II, III, and IV sites by replacing G with T residues. b, binding reactions were carried out with AR42J protein and equimolar amounts of 32P-labeled oligonucleotides in the presence (+) or absence (−) of a 100-fold molar excess of Sp1 oligonucleotide. c, binding reactions containing AR42J protein were done with 32P-labeled Sp1 oligonucleotide in the presence of cold, double-stranded competitor oligonucleotides at the concentration indicated. d, the effect of point mutations described in a upon binding was studied in reactions containing AR42J protein and 32P-labeled oligonucleotides II, III, and IV in the presence (+) or the absence (−) of a 100-fold molar excess of competing wild type (wt) or point mutant (PM) sequence. e, binding reactions with purified human Sp1 protein were carried out as described in b except that a 200-fold molar excess of cold Sp1 oligonucleotide was used for competition. The amount of total protein-DNA complex generated in binding reactions of c and e was determined by phosphorimage analysis.
In this paper we report the cloning of the gene that encodes the exocrine pancreas-specific bHLH protein Ptf1-p48 and analyze requirements for its expression. Constitutive expression of this regulatory gene, which in the animal is under tight cell-specific control, is governed exclusively by ubiquitous factors, such as Sp1 and cell-specific control, is governed exclusively by ubiquitous factors.

The Sp1 and αCbf binding sites apparently act in concert to impose a full response upon the ptf1-p48 gene promoter since none of these sites is individually capable of sustaining wild type levels of expression. In contrast, inactivation of Sp1 binding site I does not negatively affect expression of the reporter gene. Qualitatively similar results were obtained when hybrid genes were expressed in NIH3T3 cells (data not shown). As a general rule, binding sites that exhibit a high affinity for nuclear proteins from AR42J or AR41P cells were bound, either in the presence or absence of a 200-fold molar excess of cold, homologous, and heterologous competitor DNAs.

The Sp1 and αCbf binding sites apparently act in concert to impose a full response upon the ptf1-p48 gene promoter since none of these sites is individually capable of sustaining wild type levels of expression. In contrast, inactivation of Sp1 binding site I does not negatively affect expression of the reporter gene. Qualitatively similar results were obtained when hybrid genes were expressed in NIH3T3 cells (data not shown). As a general rule, binding sites that exhibit a high affinity for nuclear proteins from AR42J or AR41P cells were bound, either in the presence or absence of a 200-fold molar excess of cold, homologous, and heterologous competitor DNAs.

The Sp1 and αCbf binding sites apparently act in concert to impose a full response upon the ptf1-p48 gene promoter since none of these sites is individually capable of sustaining wild type levels of expression. In contrast, inactivation of Sp1 binding site I does not negatively affect expression of the reporter gene. Qualitatively similar results were obtained when hybrid genes were expressed in NIH3T3 cells (data not shown). As a general rule, binding sites that exhibit a high affinity for nuclear proteins from AR42J or AR41P cells were bound, either in the presence or absence of a 200-fold molar excess of cold, homologous, and heterologous competitor DNAs.

The Sp1 and αCbf binding sites apparently act in concert to impose a full response upon the ptf1-p48 gene promoter since none of these sites is individually capable of sustaining wild type levels of expression. In contrast, inactivation of Sp1 binding site I does not negatively affect expression of the reporter gene. Qualitatively similar results were obtained when hybrid genes were expressed in NIH3T3 cells (data not shown). As a general rule, binding sites that exhibit a high affinity for nuclear proteins from AR42J or AR41P cells were bound, either in the presence or absence of a 200-fold molar excess of cold, homologous, and heterologous competitor DNAs.

The Sp1 and αCbf binding sites apparently act in concert to impose a full response upon the ptf1-p48 gene promoter since none of these sites is individually capable of sustaining wild type levels of expression. In contrast, inactivation of Sp1 binding site I does not negatively affect expression of the reporter gene. Qualitatively similar results were obtained when hybrid genes were expressed in NIH3T3 cells (data not shown). As a general rule, binding sites that exhibit a high affinity for nuclear proteins from AR42J or AR41P cells were bound, either in the presence or absence of a 200-fold molar excess of cold, homologous, and heterologous competitor DNAs.

The Sp1 and αCbf binding sites apparently act in concert to impose a full response upon the ptf1-p48 gene promoter since none of these sites is individually capable of sustaining wild type levels of expression. In contrast, inactivation of Sp1 binding site I does not negatively affect expression of the reporter gene. Qualitatively similar results were obtained when hybrid genes were expressed in NIH3T3 cells (data not shown). As a general rule, binding sites that exhibit a high affinity for nuclear proteins from AR42J or AR41P cells were bound, either in the presence or absence of a 200-fold molar excess of cold, homologous, and heterologous competitor DNAs.
Transcription Control of the ptf1-p48 Subunit Gene

Fig. 8. Expression analysis of hybrid genes bearing mutated binding sites for nuclear proteins of AR42J cells. a, schematic representation of hybrid genes. The wild type (wt) construct is hybrid gene D. Its origin and that of hybrid gene E are described in the legend of Fig. 4a. Open boxes I-IV designate binding sites for nuclear proteins as deduced from footprint analysis (see Fig. 5) and encompass the DNA sequences of oligonucleotides designed for EMSA. Hybrid genes PM I-IV carry individual point mutations in binding sites I-IV, respectively. The mutated sequences are depicted as black boxes labeled PM for each construct and are those shown in Figs. 6 and 7. The relative activity of various hybrid genes is indicated by the numbers on the right and is based on data shown in b. The expression of hybrid genes was monitored by measuring enzymatic activity produced by the luciferase reporter gene. Luciferase activities were determined in two independent transfection experiments (I and II) and normalized as described in the legend of Fig. 4.

cis-acting DNA element (enhancer) that was missing in our hybrid gene constructs. To test the validity of such a model, it would have to be ultimately shown that sequences located between −900 and −3600 are accessible for the binding of such a (putative) repressor activity in the endogenous gene, and, conversely, that deletion of these sequences activates transcription in other cell types of the organism. We consider it rather unlikely, however, that the repression observed is a meaningful phenomenon, i.e., the mechanism by which cell-specific expression of the endogenous gene is assured, since it only partially inhibits expression of transgenes in cultured cells. It can also be argued that our analysis has missed cell-specific factors that do not interact with specific DNA but rather make protein-protein contacts with general DNA-binding proteins and/or cell-specific modification(s) which alter the specificity of transcription factors Sp1 or Cbf. Although such parameters might confer cell specificity to a gene promoter that is under control of general transcription factors, they do not readily explain why transgenes are active also in nonproducing cells.

While our study establishes the molecular basis for constitutive expression of the ptf1-p48 gene in cultured cells, it falls short of pinpointing the mechanism that leads to cell-specific activation of the gene in the animal. In the absence of any experimental evidence, we may only speculate how this might be achieved. The fact that we have not been able to identify a cell-specific activator in differentiated exocrine pancreatic cells in culture does not necessarily preclude that such a molecule exists at a particular stage of pancreas development. Any mechanism leading to cell-specific expression of the endogenous ptf1-p48 gene must activate the gene in the correct cell type and maintain it in a repressed state in cells of other origin. One way by which a cell- or cell lineage-specific activator might establish a cell-specific expression pattern is by converting the ptf1-p48 gene locus in the pancreatic precursor cell from an inactive into an active state, for instance by inducing cell-specific changes in chromatin structure (29) and/or DNA methylation (30). In case this factor would be a sequence-specific DNA-binding protein, its binding site might only function at a particular developmental period and not act as a transcriptional response element at later stages, including differentiated cells in culture. The altered state of the gene would then be permissive for the binding of transcriptional activators which at this stage may but do not necessarily have to be the same as those (Sp1, Cbf) ensuring maintenance of expression at the differentiated state. If such a mechanism applies, it would furnish an explanation for the apparent paradox that transgenes are not expressed in a cell type-specific fashion. It is generally accepted that foreign DNA, when introduced into cells in culture, retains a methylation status and adopts a chromatin structure permissive for expression. Therefore, if a transgene requires exclusively general transcription factors for its expression, it will be expressed regardless of the origin of the cell since the specific control mechanisms for its cell-specific activation and/or repression during embryonic development are expected to be missing in cells in culture.

There is some evidence that such a gene activation model may be valid for other regulatory genes. For instance, muscle-specific expression of the myoD gene is regulated by a distal enhancer element located 20 kb upstream of the transcription initiation site (31). However, this enhancer does not bind muscle-specific nuclear factors and confers expression to transgenes also in nonproducing cells (32). It has been suggested, therefore, that the methylation status and/or local chromatin structure govern the accessibility of this enhancer to trans-acting factors at different stages of development. A similar situation applies to the tie gene which is normally expressed in endothelial cells only (33). No cell-specific expression occurs with a transgene containing a reporter gene under control of the tie gene promoter upon transfection into cell lines in culture. However, the promoter induces correct temporal and spatial expression in transgenic mice suggesting that it contains a structural element that functions only during a particular period of embryonic life.

What might the factor that ensures correct activation in time and space of the gene encoding Ptf1-p48 during development be? Potential candidates are DNA-binding proteins that play a critical role in determining cell fate but may be only transiently expressed during embryonic development. These include, for example, homeotic factors (34) and members of the Hnf3 (forkhead) family of proteins (35–38). In this context, it is interesting to note that a homeobox protein, Msx1, represses transcription of the myoD gene in non-muscle cells by a yet not characterized interaction with its enhancer (39). We have recently shown that members of the Hnf3 class of proteins, Hnf3β and -γ, which are specifically expressed in cells of endodermal origin, are also present in the nucleus of exocrine pancreatic...
cells where they are required for the expression of the a-amylase 2 gene (40). There is no suggestive evidence to support a role of Hnf3 proteins during activation of the ptf1-p48 gene at the present time, since no binding sites occur within the established sequence (+1 to −650) of the ptf1-p48 5′-flanking region. However, we have identified within this region of the DNA a sequence element that constitutes in vitro a high affinity binding site for homeobox protein Ipf1.2 This protein, which was originally identified as a transcriptional activator of the mouse insulin gene in endocrine pancreatic β cells (41), plays a master role during pancreas ontogeny since inactivation of its gene in the animal abolishes formation of all pancreatic structures (42). The presence of this protein has been detected in the pancreatic primordium as early as day 9 of gestation (41). Its synthesis thus precedes the onset of p48 mRNA synthesis at day 12 of embryonic development (9). Even though differentiated exocrine pancreatic cells do not synthesize Ipf1, the protein is transiently expressed in cells of the exocrine lineage during development (43). This observation would be compatible with a function of Ipf1 during the process that leads to activation of the ptf1-p48 gene. It will be essential to carry out an expression analysis of hybrid genes in transgenic mice, not only to determine whether the binding site for this particular protein is indeed required for cell-specific activation of the ptf1-p48 gene promoter, but also to identify other potential DNA elements crucially involved in this process.

Acknowledgments—We thank Dr. N. Mermod for a generous gift of anti-NF1 antibody. We also thank N. Zoerkler and C. Berney for expert technical assistance, and K. Wingate for typing the manuscript.

REFERENCES

1. Braun, T., Buschhausen, D. G., Bober, E., Tannich, E., and Arnold, H. H. (1989) EMBO J. 8, 701–709
2. Davis, R. L., Weintraub, H., and Lassar, A. B. (1987) Cell 51, 987–1000
3. Edmonson, D. G., and Ohlson, E. N. (1989) Genes Dev. 3, 628–640
4. Weintraub, H. (1993) Cell 75, 1241–1244
5. Braun, T., Bober, E., Buschhausen-Denker, G., Köhtz, S., Grzeschik, K., and Arnold, H. H. (1989) EMBO J. 8, 3617–3625
6. Edmonson, D. G., Cheng, T. C., Cserjesi, P., Chakraborty, T., and Ohlson, E. N. (1992) Mol. Cell. Biol. 12, 3665–3677
7. Thayer, M. J., Tapscott, S. J., Davis, R. L., Wright, W. E., Lassar, A. B., and Weintraub, H. (1989) Cell 58, 241–248
8. Cockell, M., Stevenson, B. J., Strubin, M., Hagenbüchle, O., and Wellauer, P. K. (1989) Mol. Cell. Biol. 9, 2464–2476
9. Krapp, A., Knöfler, M., Frutiger, S., Hughes, G. H., Hagenbüchle, O., and Wellauer, P. K. (1996) EMBO J. 15, in press
10. Petrucco, S., Wellauer, P. K., and Hagenbüchle, O. (1990) Mol. Cell. Biol. 10, 254–264
11. Sommer, L., Hagenbüchle, O., Wellauer, P. K., and Strubin, M. (1991) Cell 67, 987–994
12. Roux, E., Strubin, M., Hagenbüchle, O., and Wellauer, P. K. (1989) Genes Dev. 3, 1613–1624
13. Johnson, D. A., Gauchter, J. W., Sportsman, J. R., and Elder, J. H. (1984) Gene Anal. Tech. 1, 3
14. Wellauer, P. K., and Dawid, I. B. (1977) Cell 10, 193–212
15. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463–5467
16. Hagenbüchle, O., and Wellauer, P. K. (1992) Nucleic Acids Res. 20, 3555–3559
17. Schibler, U., Hagenbüchle, O., Wellauer, P. K., and Pittet, A. C. (1983) Cell 33, 501–508
18. Hagenbüchle, O., Wellauer, P. K., Cribbs, D. L., and Schibler, U. (1984) Cell 38, 737–744
19. Samboiro, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
20. Dignam, J. D., Lebowitz, R. M., and Roeder, R. G. (1983) Nucleic Acids Res. 11, 1475–1489
21. Jessop, N. W., and Hay, R. J. (1980) In Vitro 16, 212
22. Sawadogo, M., and Sentenac, A. (1990) Nuc. Acids Res. 18, 711–715
23. Dynan, W. S., and Tjian, R. (1983) Cell 33, 75–87
24. Kadonaga, J. T., Carner, K. R., Masiasz, F. R., and Tjian, R. (1987) Cell 51, 1079–1090
25. Jones, K. A., Kadonaga, J. T., Luciwi, P. A., and Tjian, R. (1986) Science 232, 755–759
26. Jones, K. A., Kadonaga, J. T., Rosenfeld, P. J., Kiley, T. J., and Tjian, R. (1987) Cell 48, 79–89
27. Leegwater, P. A. J., van Driel, W., and van der Vliet, P. C. (1985) EMBO J. 4, 1515–1521
28. Chodosh, L. A., Baldwin, A. S., Carthew, R. W., and Sharp, P. A. (1988) Cell 53, 11–24
29. Groudine, M., and Weintraub, H. (1982) Cell 30, 131–139
30. Razin, A., and Cedar, H. (1991) Microbiol. Rev. 55, 451–458
31. Goldhammer, D. J., Bruk, S. P., Faerman, A., King, A., Shani, M., and Emerson, C. P., Jr. (1995) Development 121, 637–649
32. Goldhammer, D. J., Faerman, A., Shani, M., and Emerson, C. P., Jr. (1992) Science 256, 537–543
33. Korhonen, J., Lahtinen, I., Halmekytö, M., Alhonen, L., Jäntti, J., Dumont, D., and Alitalo, K. (1995) Blood 86, 1828–1835
34. Marata, G. (1995) Curr. Opin. Genet. Dev. 5, 606–614
35. Cleland, D. E., Overdier, D. G., Tao, W., Qian, X., Panz, L., Lai, E., and Costa, R. H. (1993) Proc. Natl. Acad. Sci. U.S.A. 90, 3948–3952
36. Lai, E., Prezioso, V. R., Tao, W., Chen, W. S., and Darnell, J. E., Jr. (1993) Genes Dev. 7, 416–427
37. McPherson, C. E., Shim, E.-Y., Friedman, D. S., and Zaret, K. S. (1993) Cell 75, 387–398
38. Szakács, G., and Hogan, B. L. (1993) Development 118, 47–59
39. Woloshin, P., Song, K., Degnin, C., McNeill Killary, A., Goldhammer, D. J., Sassoon, D., and Thayer, M. J. (1995) Cell 82, 611–620
40. Cockell, M., Stolarczyk, D., Frütiger, S., Hughes, G. H., Hagenbüchle, O., and Wellauer, P. K. (1995) Mol. Cell. Biol. 15, 2333–2341
41. Ohlson, H., Karlsson, K., and Edlund, T. (1993) EMBO J. 12, 4251–4259
42. Jonsson, J. C., Carlson, L., Edlund, T., and Edlund, H. (1994) Nature 371, 506–509
43. Guz, Y., Montminy, M. R., Stein, R., Leonard, J., Gainer, L. W., Wright, C. V. E., and Teitelman, G. (1995) Development 121, 11–18
44. Maxam, A. M., and Gilbert, W. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 560–564
45. Ben-Hattar, J., Beard, P., and Jiricny, J. (1989) Nucleic Acids Res. 17, 10179–10190
46. Kennedy, G. C., Andersen, B., and Nilson, J. H. (1990) J. Biol. Chem. 265, 6279–6285
47. Vinson, C. R., Sigler, P. B., and McKnight, S. L. (1989) Science 246, 911–916
Constitutive Expression of the Gene for the Cell-specific p48 DNA-binding Subunit of Pancreas Transcription Factor 1 in Cultured Cells Is under Control of Binding Sites for Transcription Factors Sp1 and αCbf

Knöfler Martin, Andrea Krapp, Otto Hagenbüchle and Peter K. Wellauer

J. Biol. Chem. 1996, 271:21993-22002.
doi: 10.1074/jbc.271.36.21993

Access the most updated version of this article at http://www.jbc.org/content/271/36/21993

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 45 references, 19 of which can be accessed free at http://www.jbc.org/content/271/36/21993.full.html#ref-list-1