Pan: a transcriptional regulator that binds chymotrypsin, insulin, and AP-4 enhancer motifs

Christian Nelson, Lu-Ping Shen, Andreas Meister, Eric Fodor, and William J. Rutter

Hormone Research Institute and Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143-0534 USA

Two cDNAs were cloned that encode proteins (Pan-1 and Pan-2, with predicted molecular masses of 67.4 and 67.7 kD, respectively) that bind to a conserved transcriptional regulatory element present in pancreatic exocrine genes. The cDNAs are virtually identical in the region that encodes the amino-terminal 525 amino acids; they differ only by a 3-nucleotide insertion that results in the addition of a single amino acid. The cDNAs, however, code for related, but divergent, carboxy-terminal regions. Expression of the cloned cDNAs in HeLa cells activates expression of a cotransfected chimeric gene containing multimers of the pancreatic conserved element linked to a minimal promoter. Pan-1 and Pan-2 bind to related motifs present in the rat insulin enhancer as well as a DNA segment containing the SV40 AP-4 element. The Pan-1 and Pan-2 cDNAs display extensive sequence identity with partial cDNA clones encoding El2 and E47, which bind to the immunoglobulin κE2 cis-active element, and may be derived from the same gene. These factors belong to a set of genetically distinct molecules, including AP-4 and MLTF, that bind to the CACCTGTC motif or related sequences. The Pan/E12,E47 proteins also show structural similarity with the Drosophila daughterless protein, MyoD, Myogenin, and Myf-5.

[Key Words: Acinar cells, pancreas, cell-specific expression; DNA-binding protein]

Received August 25, 1989; revised version accepted March 16, 1990.

Exocrine cells of the pancreas synthesize a set of hydrolytic enzymes involved in the digestion of food. The dramatic increase in the levels of these enzymes during pancreatic development is due primarily to an increased rate of transcription (Rutter et al. 1978). Transcriptional enhancers have been identified that appear to determine acinar-specific expression of the various exocrine genes in both cell culture and transgenic animals. A conserved sequence has been identified in the enhancers of the chymotrypsin B, amylase 2A, trypsin I and elastase I exocrine genes, and the 5'-flanking sequences of other exocrine genes. This pancreatic consensus sequence (Swift et al. 1984; Boulet et al. 1986) has been shown to be a critical component of the functional enhancer (Boulet et al. 1986; Kruse et al. 1988).

A pancreas-specific DNase I footprint encompassing the consensus sequence has been identified in the chymotrypsin B, amylase 2A, and trypsin I genes (Meister et al. 1989). Recently, Cockell et al. (1989) identified an apparent pancreas-specific binding activity that interacts with the consensus sequence of the α-amylase 2, elastase 2, trypsin, and other pancreatic genes. Competition for binding in gel mobility-shift experiments and methylation interference studies suggest that this chymotrypsin enhancer element contains two distinct binding motifs. One motif contains the sequence, CACCTGTC, and the other, TTTTCCC (TC motif) (Cockell et al. 1989; Meister et al. 1989). These sequences form an apparent bipartite binding motif that is also present in the consensus sequence of other pancreatic genes (Cockell et al. 1989). Multimers of the chymotrypsin B sequence can function as a cell-specific enhancer in gene transfer experiments [Meister et al. 1989]. Deletion of sequences containing these elements in the rat chymotrypsin B enhancer virtually abolishes enhancer function [Boulet et al. 1986]. A mutation of the related CAGCTGTG motif present in the rat amylase 2A enhancer destroys enhancer activity [Boulet et al. 1986]. An apparent pancreas-specific binding activity protects the sequence CACCTGTC in the rat elastase I enhancer from DNase I digestion; mutations of this sequence severely decrease elastase gene expression [Kruse et al. 1988]. These collective studies suggest a crucial role for this element in the control of pancreas-specific gene expression.

Here, we report the isolation of two cDNAs that encode proteins that selectively bind to the CACCTGTC motif and demonstrate that expression of these cDNAs
in cotransfection experiments activates the expression of a gene containing multimers of the CACCTGTC motif linked to a minimal promoter.

Results

Isolation and characterization of cDNA clones

To study the mechanism of transcriptional regulation, we screened a λgt11 cDNA expression library prepared from an acinar cell line for cDNAs that encode proteins binding to the pancreatic consensus element. Of 400,000 plaques screened with a multimerized consensus sequence probe, one recombinant phage was isolated that displayed specific binding activity. A second cDNA insert that encodes an apparent full-length protein was obtained by oligonucleotide hybridization screening of the λgt11 library. The proteins designated Pan-1 and Pan-2, encoded by the cDNA inserts were examined for their binding specificity. Pan-1, produced in vitro, displayed selective binding to a region containing the CACCTGTC motif in methylation interference and binding competition experiments [Fig. 1A,B]. Pan-2 produced by in vitro translation also selectively binds the pancreatic consensus sequence. This binding is abolished by a block mutation of the CACCTGTC motif [see Fig. 6, below]. Deletion of the DNA sequences coding for the carboxy-terminal 224 amino acids of Pan-1 abolished the binding ability of in vitro-translated protein, whereas removal of coding sequences for the 23 carboxy-terminal amino acids only decreased the binding affinity [data not shown]. Thus, the carboxy-terminal region may be involved in DNA binding.

The 2.7-kb cDNA insert encoding Pan-1 and the 3.0-kb insert encoding Pan-2 were sequenced and found to be identical except for a region at the 3' terminus and a 3-nucleotide insertion in the Pan-1 cDNA. Therefore, the two cDNAs are probably derived from the same gene. The Pan-1 cDNA contains an open reading frame coding for 638 amino acids [Fig. 2]. The Pan-2 cDNA insert contains an open reading frame coding for an apparent full-length protein of 649 amino acids [with a predicted molecular mass of 67.7 kD]. The Pan-2 cDNA contains the coding sequences for the initiator methionine and the first nine amino acids, which are absent in the Pan-1 cDNA. With the exception of an additional alanine present at predicted amino acid residue 168 in Pan-1, the two proteins diverge only in a 72-amino-acid region of the carboxy-terminal section (amino acids 525-597) of the carboxy-terminal region. They regain identity from amino acid 597 to the divergence at 525 follows a CAG consensus splice donor sequence (Mount 1982). Thus, the two cDNAs may be generated by alternative splicing. The Pan-1 and Pan-2 cDNAs possess extensive nucleotide similarity in the divergent region, suggesting that alternative splicing of an exon duplication is responsible for the unique cDNAs. A DNA sequence present in both Pan-1 and Pan-2 cDNAs hybridizes to single BamHI and HindIII genomic restriction fragments [Fig. 3]. This result supports the premise that a single gene in the rat codes for the two proteins.

Interestingly, Pan-1 and Pan-2 are very closely related to E12 and E47, factors that bind to the immunoglobulin κE2 motif and that have been characterized by partial cDNA clones isolated from a human cDNA library [Murre et al. 1989]. Pan-1 corresponds to E47 [91% identity over 131 amino acids] and Pan-2 to the E12 clone [83% identity over 440 amino acids]. E12 and E47 diverge at the same position as Pan-1 and Pan-2 in the carboxy-terminal region. Therefore, the clones may be derived from a single gene in the rat.
derived from the same gene. The differences in amino acid composition between Pan-1/2 and E12/47 may be due to evolutionary changes in the rat and human forms of the same proteins. In addition, these proteins are closely re-

Figure 2. [See following page for remainder of figure and for legend.]

GENES & DEVELOPMENT 1037
Figure 2. Nucleotide sequence of Pan-1 and Pan-2 cDNAs and predicted amino acid sequences. An in-frame TAG termination codon 5′ of the presumed initiator methionine of Pan-2 is underlined. A linkered EcoRI site (GAATTCC) at the 5′ end of the Pan-1 cDNA is also underlined. The nucleotide sequences of the Pan-1 and Pan-2 cDNAs are identical except for a GCA insertion that codes for an Ala residue at amino acid 168 of Pan-1 and a carboxy-terminal divergent region extending from amino acid 525 to 597 of Pan-2. The Leu components of a leucine zipper motif are circled. Sequencing of DNA was performed by use of the Sanger dideoxynucleotide method.

related to the Drosophila daughterless protein (Caudy et al. 1988) and share a more limited structural similarity to c-myc (Battey et al. 1983) and proteins implicated in muscle cell differentiation, such as MyoD (Davis et al. 1987), Myogenin (Wright et al. 1989), and Myf-5 (Braun et al. 1989) in the putative DNA-binding domain (Fig. 4) (Murre et al. 1989). This region has been postulated to form an amphipathic helix−loop−helix involved in protein dimerization (Murre et al. 1989).

The protein sequences are distinctive in various regions of the molecule, suggesting the presence of functional domains. Five methionines are clustered at the amino-terminus (Met-Met-5-Met-16-Met-Met). With the exception of the first methionine, all amino-terminal methionines obey the Kozak rules for efficient initiation of translation (Kozak 1989). Thus, they may serve as alternative sites for the initiation of translation. At the amino terminus, there is a short potential helical structure (which would be eliminated if translation initiated at position 25). The next 467 amino acids are rich in proline (11%) and glycine (14%) residues, thus it is unlikely that there are any extended helical motifs. This region is extraordinarily rich in hydroxy amino acids [serine 18%, tyrosine 3%] that represent potential phosphorylation sites. Tyrosine residues are present only in the region from amino acid 76 to 378 and are especially concentrated [16%) from amino acid 186 to 216. A putative leucine zipper motif (Landschulz et al. 1988) is found in the

Figure 3. Southern blot analysis and restriction map of the Pan gene. (Top: A, B, C, and D) Rat genomic DNA was digested with EcoRI (lane 1), BamHI (lane 2), HindIII (lane 3), EcoRI and BamHI (lane 4), EcoRI and HindIII (lane 5), and HindIII and BamHI (lane 6). (E) Rat genomic DNA was digested with EcoRI (lane 1), BamHI (lane 2), HindIII (lane 3), HindIII and BamHI (lane 4), EcoRI and HindIII (lane 5), and EcoRI and BamHI (lane 6). The transferred DNA was probed with: [A] HindIII−BamHI fragment from the Pan-1 cDNA coding for predicted amino acids 424−627 and encompassing the 3′ divergent region, [B] the HindIII−BamHI fragment from the Pan-2 cDNA coding for predicted amino acids 423−626 and encompassing the 3′ divergent region, [C] the entire Pan-1 cDNA, [D] the entire Pan-2 cDNA, and [E] EcoRI−PstI fragment from the Pan-1 cDNA coding for predicted amino acids 10−163. (Bottom) Restriction map of the Pan gene.
region from amino acid 385 to 420 [Leu(385)-Leu(392)-
Leu(399)-Thr(406)-Leu(413)-Leu(420)]. In this region, similar sequences {Leu/Ile-His-Gly/Val-Leu} are found ending at alternating heptads at leucines 385, 399, and 413. If this region of the protein forms a coiled-coil structure, as is proposed for leucine zipper dimers, histidine residues 383, 397, and 411 would align on one face of the helix. The region surrounding and including the leucine zipper motif (276–457) is rich in histidine residues (6%). The region from amino acid 383 to 457 is especially histidine rich (12%). Two copies of a Ser-Ser-Ser-Gly-Asp repeat are found at positions 188 and 322. This peptide is also found in the human lamin A protein, which shares a limited structural similarity to the DNA-binding region of the El2 and E47 proteins (Fisher et al. 1986; Murre et al. 1989). A second repeat [Pro-Ser-Ser-Val/Ala-Tyr-Pro-Ser] is present at positions 182 and 207. The carboxy-terminal region beginning at amino acid 492 contains several features that distinguish it from the rest of the molecule, including periodic highly charged domains, a strongly hydrophobic domain, and the potential to form extended α helices. The region in the α2 binding proteins equivalent to amino acids 515–605 of Pan-2 contains the DNA-binding and dimerization domains [Murre et al. 1989]. The carboxy-terminal region contains no tyrosine residues, and serines decrease from 18% to 4%. The abundance of glycine decreases from 12% in the first 490 residues to less than 1%; only 1 glycine is present in the region from amino acid 490 to 623. Highly charged domains in Pan-1 and Pan-2, comprised of both positive and negative charges, begin at positions 493, 508, 536, 551, and 614. Pan-2 contains an additional charged motif (comprised of 6 consecutive negatively charged amino acids) beginning at residue 525. A hydrophobic stretch of amino acids, present from 584 to 597, could form a short α-helix that corresponds to the amphipathic helix II postulated to form the helix–loop–helix structure of the E12- and E47-encoded proteins [Murre et al. 1989].

Expression of Pan cDNAs activates expression of a promoter containing multimers of the CACCTGTC motif

The putative transcriptional regulatory function of Pan-1 and Pan-2 was examined by a test of the ability of these proteins to regulate the expression of a reporter gene comprised of seven repeats of the chymotrypsin consensus sequence linked to a prolactin minipromoter that drives transcription of the firefly luciferase gene [Chy7–Prl–Luc] (Nelson et al. 1988). This fusion gene is expressed at high levels in the cell lines AR4-2 and 266-6, which synthesize exocrine gene products, but is silent in HeLa cells [Meister et al. 1989]. The Pan-2 cDNA driven by a cytomegalovirus enhancer and promoter selectively activates expression of the Ch7–Prl–Luc fusion gene when cotransfected into HeLa cells (Fig. 5). The same effect was observed when the carboxy-terminal region of Pan-2 was replaced with the similar region of Pan-1 in the expression vector. In contrast, cotransfection of a fusion gene encoding a truncated Pan-2 protein lacking 226 amino acids of the carboxyl terminus was unable to activate expression (Fig. 5). Thus, Pan-1 and Pan-2 can apparently function as transcriptional activators.

Pan-1 and Pan-2 bind insulin and SV40 enhancer motifs

Because regulatory elements of the rat insulin I gene and SV40 enhancer contain sequences related to the CACCTGTC motif of the pancreatic consensus sequence (Table 1), we tested whether the Pan-1 and Pan-2 proteins could bind to these elements in gel mobility-shift experiments. In vitro-translated Pan-1 and Pan-2 bind to the rat insulin I "Nir box" [Moss et al. 1988] as well as to a 38-bp region of the SV40 virus that contains the high affinity AP-4 binding element [Mermod et al. 1988] (Fig. 6). Pan-1 binds with a higher affinity than Pan-2 to the chymotrypsin B, Nir, and SV40 elements. Neither protein was able to bind to the chymotrypsin B element containing a mutated CACCTGTC motif. A comparison of the sequences bound by Pan-1 and Pan-2 suggests a binding site comprised of the consensus sequence of CAc/cCTGT/c (Table 1). This CACCTGTC motif is present in the amylase 2A, chymotrypsin B, elastase I and trypsin 1 enhancers, and the 5'-flanking region of other genes expressed in acinar cells of the pancreas [Boulet et al. 1986; Cockell et al. 1989].

Discussion

cDNAs have been isolated independently from an HIT pancreatic endocrine cell line library by expression screening with an insulin Nir element oligomer, which encodes proteins that are highly related to Pan, E12, and

Figure 4. Related amino acids among Pan-1, Pan-2, the Drosophila daughterless protein [Caudy et al. 1988], and proteins implicated in muscle differentiation: MyoD [Davis et al. 1987], Myogenin [Wright et al. 1989], and Myf-5 [Braun et al. 1989]. The boxed region represents identical amino acids. The regions compared are amino acids 535–607 of Pan-2, 533–605 of Pan-1, 545–617 of daughterless, 75–145 of Myf-5, 100–170 of MyoD, and 72–142 of Myogenin. The protein sequences were aligned using the Doolittle sequence analysis program in the Eu Gene package.
E47. The differences detected are likely to be due to alternative splicing (M. German et al., in prep.). These cDNAs are likely to be derived from the same gene, which encodes Pan-1 and Pan-2. Because a single 3.0-kb transcript is detected in all rat tissues tested (Fig. 7), and cDNAs encoding Pan/E12- and E47-related proteins isolated from lymphoid and pancreatic exocrine and endocrine cell lines, it is likely that Pan/E12,E47 gene products are distributed widely. In addition, apparent binding sites for Pan-1 and Pan-2 are also distributed widely. Other factors also cross-react with the CACCTGTC or related motifs, thereby increasing the binding repertoire and the opportunity for control over the transcriptional regulatory system. Purified AP-4 binds to the CACCTGTC motif present in the amylase 2A and other exocrine gene enhancers (Fodor et al., in prep.) and, conversely, Pan-1 and Pan-2 bind to the AP-4 element in SV40. The adenovirus MLTF likely binds to the CACCTGTC element because the MLTF binding site competes with the chymotrypsin B and insulin N' elements in binding experiments using pancreatic exocrine and endocrine nuclear extracts (Moss et al. 1988; Meister et al. 1989). However, both the AP-4 gene (R. Tjian, pers. comm.) and the MLTF gene (R. Roeder, pers. comm.) appear to be distinct from the Pan/E12,E47 gene.

Functional specificity therefore may be determined by specific protein–protein interactions. For example, Pan proteins may interact with other factors that bind to the CACCTGTC element or to the contiguous TTTCCT motif. The putative leucine zipper, or carboxy-terminal dimerization domains may provide the structural basis for dimerization as in the case of Fos and Jun (Kouzarides and Ziff 1988). Such dimerization may be the fundamental unit of the activation complex, perhaps by providing a common structural substrate for some other functional (linking?) protein. The affinities and specificities of the various interacting factors then may determine the functional competence of the complex. The understanding of the structure and function of this fundamental unit may lead to an understanding of the mechanisms underlying specific gene expression.

Materials and methods

Isolation of cDNA clones

An oligo(dT)-primed library was constructed using poly(A)-selected RNA prepared from the rat exocrine cell line AR4-2J. Nick-translated DNA comprised of multimerized consensus sequence was used to probe the library according to Vinson et al. (1988). The probe was composed of two complementary oligonucleotides of the following sequence, which includes rat chymotrypsin gene sequences from -213 to -186:

5'-GATCTCAGGGCACCTGTCCTTTTCCCATGGCCGAGTCCCGTGGACAGGAAAAGGGTACCGGCCTAG-3' 

The Pan-2 cDNA was isolated by hybridization screening of the AR4-2J library with an oligonucleotide complementary to a sequence present in the 5' region of Pan-1 cDNA. The oligonucleotide is comprised of the following sequence:

5'-TAGCGGGAACATCATACTGAAGTCCAGGAGGTC-3'

Plasmid constructions

Plasmids used as templates for in vitro transcription were constructed by cloning Pan-1 and Pan-2 cDNAs into the EcoRI site of a pGEM-3Z vector (Promega). Expression vector plasmids were constructed by inserting Pan cDNA sequences into the cytomegalovirus enhancer- and promoter-containing expression vector pLPAAS (gift of M. Walters, unpubl.). A Stul site 44 nucleotides 5' of the predicted Pan-2 initiator methionine

Figure 5. Sequence-dependent activation of gene expression by Pan proteins. Reporter plasmid (10 µg) with [Chy7Prl] or without [Prl] seven copies of the chymotrypsin pancreatic consensus element (-213 to -186) linked to a prolactin minipromoter luciferase fusion gene (Nelson et al. 1988) was cotransfected with 2 µg of Pan expression plasmids into HeLa cells. Cells were harvested 48 hr later, and luciferase activity was measured. Expression of Pan A, Pan-1, and Pan-2 cDNAs was controlled by a cytomegalovirus enhancer- and promoter-containing expression vector pLPAAS (gift of M. Walters, unpubl.). A Stul site 44 nucleotides 5’ of the predicted Pan-2 initiator methionine
Table 1. Sequence comparison of Pan binding elements and related enhancer motifs

| Species   | Enhancer   | Sequence       | Element |
|-----------|------------|----------------|---------|
| Rat       | chymotrypsin B | CACCTGT       | Pan     |
| Rat       | amylase 2A | CAGCTGT       | Pan     |
| Rat       | elastase I | CACCTGT       | Pan     |
| Rat       | trypsin I | CACCTGT       | Pan     |
| Rat       | carboxy A1 | CACCTGT       | Pan     |
| Human     | immunoglobulin μ heavy chain | CACCTGC       | μE2     |
| Human     | immunoglobulin κ heavy chain | CACCTGC       | κE2     |
| Simian    | simian virus 40 | CACCTGT       | AP-4    |
| Rat       | insulin 1 | CATCTGT       | Nis     |

Nucleotide sequences in the pancreatic consensus sequence present in the rat pancreatic exocrine genes: chymotrypsin B, amylase 2A, elastase I, trypsin I, and carboxypeptidase A1 (Boulet et al. 1986) are compared to the rat insulin I gene Nis enhancer element (Moss et al. 1988), immunoglobulin enhancer elements (Murre et al. 1989), and the SV40 AP-4 element (Mermod et al. 1988). A proposed Pan consensus binding site is depicted at bottom.

codon was converted to a BglII site by cleavage and ligation to BglII linkers. Pan-2 cDNA coding for a predicted full-length protein containing BglII and XbaI ends was inserted into the BamHI and XbaI sites of pLPAAS. Pan-1 was constructed by the replacement of the HindIII-Xhol fragment of Pan-2 with the HindIII-Xhol fragment of Pan-1 (Fig. 2), which resulted in the exchange of the carboxy-terminal divergent region of Pan-2 with that of Pan-1. Pan Δ was constructed by the insertion of a BglII–HindIII fragment of Pan-2 cDNA, which resulted in the deletion of sequences coding for the 226 carboxy-terminal amino acids of Pan-2. The reporter plasmid Chy7Prl contained seven copies of the rat chymotrypsin B consensus sequence (−213 to −186) linked to a prolactin minipromoter luciferase fusion gene PrlLuc (Nelson et al. 1988).

Gel mobility-shift assays

Linearized template (5 μg) was transcribed using T7 polymerase according to the manufacturer’s directions (Promega). One-third of the RNA produced was used in a 40-μl in vitro translation reaction using rabbit reticulocyte lysate as recommended by the manufacturer (Promega). The rabbit reticulocyte lysate (3 μl) was incubated with 200 pg of 32P 5' end-labeled probe (10,000 cpm) in a 6-μl final volume of a binding reaction containing 70 mM potassium acetate, 15 mM Tris (pH 7.8), 1 mM EDTA, 5% glycerol, 5 mM MgCl2, and 20–100 ng of poly[d(I-C)]. After incubating at room temperature for 15 min, the reaction mix was electrophoresed on a 5% nondenaturing polyacrylamide gel. Identical translation reactions using 35S-labeled methionine were performed in parallel with the Pan-1, Pan-2, and truncated templates to verify that equivalent amounts of protein were synthesized. The labeled translation reactions were analyzed on 10% acrylamide gels containing SDS. For binding competition experiments, 2- or 20-fold molar excess of competing unlabeled DNA was added to the labeled probe prior to the addition of lysate. The rat chymotrypsin B probe consisted of sequences extending from −213 to −186 (Boulet et al. 1986). The rat insulin Nis element probe consisted of sequences extending from −124 to −93 of the rat insulin I gene (Moss et al. 1988). The SV40 probe containing the high-affinity AP-4 binding site consisted of sequences extending from 279 to 246

Figure 6. Pan-1 and Pan-2 bind chymotrypsin, insulin, and SV40 enhancer elements. In vitro-produced Pan-1 and Pan-2 were incubated with radiolabeled enhancer element probes and binding was analyzed by gel mobility-shift assays. (Lanes a) Rabbit reticulocyte lysate containing in vitro-translated BMV control RNA (Promega) and specific enhancer probe; (lanes b) lysate containing in vitro-translated Pan-1 and specific enhancer probe; (lanes c) lysate containing in vitro-translated Pan-2 and specific enhancer probe. Lanes designated Chymo contain a chymotrypsin B consensus element probe. Lanes designated Insulin contain the rat I insulin Nis element probe. Lanes designated SV40 contain a region that includes the SV40 AP-4 site as a probe. Lanes designated B Mut contain chymotrypsin element block mutant B (Fig. 1), which has a mutation of the CACCTGTC sequence of the chymotrypsin B element as a probe.

Figure 7. Northern blot analysis of Pan transcripts in rat tissues and cell cultures. RNA prepared from rat tissues, the AR4-2J (Jessop and Hay 1980) and 266-6 (Ornitz et al. 1985) pancreatic exocrine cell lines, and the HIT pancreatic endocrine cell line (Edlund et al. 1985) were hybridized with a 32P-labeled Pan-2 restriction fragment 2.5-kb (EcoRI–BamHI fragment). A single 3.0-kb hybridizing species is detected in all samples tested.
of the SV40 viral genome (Mermod et al. 1988). All probes consisted of complementary annealed oligonucleotides 32P-end-labeled to approximately equal specific activities.

Methylation interference assays

For methylation interference experiments, a preparative binding reaction was conducted using a 10-fold scale-up of the binding conditions described. The radiolabeled probe was partially methylated (Maxam and Gilbert 1980), and the bound complex was visualized and isolated by autoradiography following native gel electrophoresis as described previously. DNA fragments were recovered by electroelution and subjected to piperidine cleavage according to Maxam and Gilbert (1980). The cleavage products were resolved on 12% polyacrylamide gels containing 8 M urea.

Cell culture and transfections

HeLa cells were maintained in Dulbecco’s modified Eagle medium supplemented with 10% fetal calf serum and 4.5 mg of glucose per ml. Transfections were performed using calcium phosphate precipitates containing 10 μg of reporter plasmid and 2 μg of expression plasmid as described (Nelson et al. 1988).

Luciferase assays

Luciferase assays were performed using 50 μg of protein extract as described [Nelson et al. 1988]. Background activity was measured by performing luciferase assays without added extract, and the background activity subtracted from each measured value.

Southern blot analysis

Rat liver genomic DNA (9 μg) was digested with 150 units of EcoRI, BamHI, or HindIII restriction enzymes, separately or in combination, and electrophoresed on a 0.8 % agarose gel. The DNA was transferred to an Optibond nitrocellulose membrane (Schleicher & Schuell). The transfer of DNA to the membrane, hybridizations, and washes were performed according to the manufacturer’s directions [Schleicher & Schuell]. The final wash was performed using a 0.1 x SSC solution containing 10 μg of reporter plasmid and 20 μg of expression plasmid as described [Nelson et al. 1988].

Northern blot analysis

Total cellular RNA was prepared according to Chirgwin et al. (1979). Total cellular RNA (15 μg) or poly[A]-selected RNA (pancreas only, 8 μg) were resolved by electrophoresis on a 1.3 % agarose gel containing formaldehyde and transferred to a GeneScreen membrane according to the manufacturer’s directions [New England Nuclear]. Then, the RNA was hybridized with 5 × 106 cpm per milliliter of 32P-radiolabeled probe. The blotting, hybridization, and wash conditions were carried out according to the manufacturer’s directions [New England Nuclear]. RNA size standards [BRL] were used to calculate the size of hybridizing species.

Acknowledgments

The authors thank Jeffrey Edman for advice on cDNA library construction, Fred Cohen and Samuel Karlin for reviewing protein sequence, Scott Weinrich and Michael Blanar for critically reading the manuscript, and Mike German and Larry Moss for helpful discussions. The work was supported by National Institutes of Health grant AM 21344 to W.J.R.

Note added in proof

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

References

Battey, J., C. Moulding, R. Taub, W. Murphy, T. Stewart, H. Potter, G. Lenoir, and P. Leder. 1983. The human c-myc oncogene: Structural consequences of translocation into the IgH locus in Burkitt lymphoma. Cell 34: 779–187.

Boulet, A.M., C.R. Erwin, and W.J. Rutter. 1986. Cell-specific enhancers in the rat exocrine pancreas. Proc. Natl. Acad. Sci. 83: 3599–3603.

Braun, T., G. Buschhausen-Denker, E. Bober, E. Tannich, and H.H. Arnold. 1989. A novel human muscle factor related to but distinct from MyoD1 induces myogenic conversion in 10T1/2 fibroblasts. EMBO J. 8: 701–709.

Caudy, M., H. Vassin, M. Brand, R. Tuma, L.Y. Jan, and Y.N. Jan. 1988. daughtersless, a gene essential for both neurogenesis and sex determination in Drosophila, has sequence similarities to myc and the achaete-scute complex. Cell 55: 1061–1067.

Chirgwin, J.M., A.E. Przybyla, R.J. MacDonald, and W.J. Rutter. 1979. Isolation of biologically active RNA from sources enriched in ribonuclease. Biochemistry 19: 5294–5299.

Cockell, M., B.J. Stevenson, M. Strubin, O. Hagenbuchle, and P.K. Wellauer. 1989. Identification of a cell-specific DNA-binding activity that interacts with a transcriptional activator of genes expressed in the acinar pancreas. Mol. Cell. Biol. 9: 2464–2476.

Davis, R.L., H. Weintraub, and H.B. Lassar. 1987. Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell 51: 987–1000.

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

Edlund, T., M.D. Walker, P.J. Barr, and W.J. Rutter. 1985. Cell-specific expression of the rat insulin gene: Evidence for role of two distinct 5’ flanking elements. Science 230: 912–916.

Fisher, D.Z., N. Chaudhary, and G. Blobel. 1986. cDNA sequencing of nuclear lamins A and C reveals primary and secondary structural homology to intermediate filament proteins. Proc. Natl. Acad. Sci. 83: 6450–6454.

Fried, M. and D.M. Crothers. 1981. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. 9: 6505–6525.

Garner, M.M. and A. Revzin. 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: Application to components of the Escherichia coli lactose operon regulating system. Nucleic Acids Res. 9: 3007–3060.

Jessop, N.W. and R.J. Hay. 1980. Characteristics of two rat pancreatic exocrine cell lines derived from transplantable tumors. In Vitro 16: 212.

Kouzarides, T. and E. Ziff. 1988. The role of the leucine zipper in the fos-jun interaction. Nature 336: 646–651.

Kozak, M. 1989. The scanning model for translocation: An update. J. Cell Biol. 108: 229–241.

Kruuse, F., C.T. Komro, C.H. Michnoff, and R.J. MacDonald. 1988. The cell-specific elastase I enhancer comprises two domains. Mol. Cell. Biol. 8: 893–902.

Landschulz, W.H., J.R. deWet, J.R., K.V. Wood, M. DeLuca, D.R. Helinski, and S. Subramani. 1987. Firefly luciferase gene: Structure and expres-
Regulation of pancreatic enhancer function

DNA with base-specific chemical cleavages. *Methods Enzymol.* 65: 499–560.

Meister, A., S.L. Weinrich, C. Nelson, and W.J. Rutter. 1989. The chymotrypsin enhancer core: Specific factor binding and biological activity. *J. Biol. Chem.* 264: 20744–20751.

Mermod, N., T.J. Williams, and R. Tjian. 1988. Enhancer binding factors AP-4 and AP-1 act in concert to activate SV40 late transcription in vitro. *Nature* 332: 557–561.

Moss, L.G., J.B. Moss, and W.J. Rutter. 1988. Systematic binding analysis of the insulin gene transcription control region: Insulin and immunoglobulin enhancers utilize similar transactivators. *Mol. Cell. Biol.* 8: 2620–2627.

Mount, S.M. 1982. A catalogue of splice junction sequences. *Nucleic Acids Res.* 10: 459–472.

Murre, C., P.S. McCaw, and D. Baltimore. 1989. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD and myc proteins. *Cell* 56: 777–783.

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

Ornitz, D.M., R.D. Palmeter, R.E. Hammer, R.L. Brinster, G.H., Swift, and R.J. MacDonald. 1985. Specific expression of an elastase-human growth hormone fusion gene in pancreatic acinar cells of transgenic mice. *Nature* 313: 600–602.

Rutter, W.J., A.E. Przybyla, R.J. MacDonald, J.D. Harding, J.M. Chirgwin, and R.L. Pictet. 1978. Pancreas development: An analysis of differentiation at the transcriptional level. In *Cell differentiation and neoplasia*, pp. 487–508. Raven Press, New York.

Swift, G.H., C.S. Craik, S.J. Stary, C. Quinto, R.G. Lahaie, W.J. Rutter, and R.J. Macdonald. 1984. Structure of the two related elastase genes expressed in the rat pancreas. *J. Biol. Chem.* 259: 14271–14278.

Vinson, C.R., K.L. LaMarco, P.F. Johnson, W.H. Landschulz, and S.L. McKnight. 1988. *In situ* detection of sequence-specific DNA binding activity specified by a recombinant bacteriophage. *Genes Dev.* 2: 801–806.

Wright, W.E., D.A. Sasoon, and V.K. Lin. 1989. Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD. *Cell* 56: 607–617.
Pan: a transcriptional regulator that binds chymotrypsin, insulin, and AP-4 enhancer motifs.

C Nelson, L P Shen, A Meister, et al.

*Genes Dev.* 1990, 4:
Access the most recent version at doi:10.1101/gad.4.6.1035

References
This article cites 28 articles, 13 of which can be accessed free at:
http://genesdev.cshlp.org/content/4/6/1035.full.html#ref-list-1

License

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