The cell-specific transcription factor PTF1 contains two different subunits that interact with the DNA

E. Roux, M. Strubin, O. Hagenbüchle, and P.K. Wellauer
Swiss Cancer Research Institute, 1066 Epalinges, Switzerland

The cognate sequence of transcription factor PTF1, which plays a key role in pancreas-specific gene expression, has a bipartite organization. Two separate DNA domains, the A and the B boxes, are required for efficient binding of the factor. The structure of PTF1 was elucidated by cross-linking purified PTF1 to DNA templates that had been differentially substituted with azido-deoxyuridine (N3 • dU). This site-directed UV cross-linking shows that PTF1 contains two DNA-binding proteins, distinct in size and sensitivity to Staphylococcus aureus V8 protease. A 64-kD protein is cross-linked with DNA containing N3 • dU substitutions in the A box, and a 48-kD protein is cross-linked with DNA containing N3 • dU substitutions in the B box. Both proteins bind simultaneously to the same DNA molecule. The data indicate that PTF1 is a heteromeric oligomer and that its cell-specific DNA-binding potential is the result of a concerted activity of two DNA-binding subunits.

[Key Words: Cell-specific transcription factor; UV cross-linking; protein–DNA interaction; bipartite recognition motif; heteromeric oligomer]

Received March 12, 1989; revised version accepted July 14, 1989.

Pancreas-specific gene expression requires the presence of the transcription factor PTF1, which binds to sequence elements located in the promoter region of genes expressed in this tissue (Cockell et al. 1989). The DNA-binding activity of PTF1 has been originally identified in rat pancreatic cells in culture but also has been detected in mouse pancreas. The activity of this factor, which is restricted to acinar pancreatic cells, is detected at all stages of development at which the promoters of genes encoding pancreas-specific functions are active (Petrucio et al., in prep.).

The interaction of PTF1 with its cognate sequence has been studied in some detail by G methylation interference and protection footprinting (Cockell et al. 1989). The sequence requirements for binding of this factor have been elucidated for a number of genes, including those encoding α-amylase, trypsin, elastase, carboxypeptidase, and chymotrypsinogen. The PTF1 recognition sites of all these genes, although degenerate in sequence, have certain features in common, the most striking of which is a bipartite organization of DNA sequences that are essential for factor binding. These sequences are arranged in two distinct clusters, the A and the B boxes, and are separated by DNA in which no interaction with the protein can be detected. The sequences necessary for binding of the factor coincide with those specifying two domains of high sequence homology in the various genes. The PTF1-binding site is subject to stereo-specific constraints because factor binding requires the two domains of protein–DNA interaction to be spaced by integral helical turns of B-form DNA. In general, the sequences that interact with the factor are separated by one turn of the DNA helix. The exception to this rule is the α-amylase gene element in which these regions are separated by two helical turns of DNA. This peculiar organization of the PTF1-binding site is compatible with models involving either a single protein or multiple proteins that bind in an interdependent manner. In each of these models, either the protein and/or the DNA must be able to assume alternative conformations.

To elucidate the structure of PTF1, we carried out UV cross-linking experiments with the α-amylase cognate sequence and purified PTF1. In this paper we show that PTF1 belongs to a novel class of eukaryotic transcription factors. It interacts with DNA through two subunits that differ in size and primary sequence and exhibit different DNA sequence specificities.

Results

PTF1 is a single DNA-binding entity

A characteristic feature of various PTF1 recognition motifs is their bipartite organization with respect to protein binding (Cockell et al. 1989). In the α-amylase gene element, the centers of the A and B box domains of protein–DNA interaction are separated by approximately two integral turns of the DNA helix (Fig. 1A). We have shown that binding reactions containing either crude...
Figure 1. The interaction of purified PTF1 with complete and partial PTF1-binding sites. (A) Map of the synthetic oligonucleotide Amy2-IV (Cockell et al. 1989). This oligonucleotide encompasses the PTF1-binding site of the α-amylase gene (nucleotides −121 to −157), as defined by footprint analysis, and is flanked by synthetic BamHI sites. The nucleotides specifying the A and B box elements, which are homologous to PTF1 motifs of other genes expressed in the pancreas, are boxed. ■ The G residues at which methylation interferes with factor binding. T residues, which are subject to substitution by N3-dU, are numbered. (B) Gel retardation analysis. The following DNA probes were used: the Amy2-IV oligonucleotide (A), which represents the complete PTF1-binding site; the Amy2-IVA oligonucleotide, which contains sequences −130 to −157 and lacks most of the B box; the Amy2-IVB oligonucleotide, which encompasses sequences −114 to −142 of the α-amylase gene and lacks the entire A box. Probes end-labeled with 32P [0.2 ng] were incubated with the fraction containing the α form of PTF1, which was obtained by chromatography on a Sephacryl S-300 column and was purified further on a DNA affinity column. DNA–protein complex was separated from unbound DNA on a 2% agarose gel. (Lanes 1, 4, and 8) Standard binding reactions containing Amy2-IV oligonucleotide; (lanes 5 and 9) standard binding reactions containing oligonucleotides Amy2-IVB and Amy2-IVA, respectively; (lanes 2, 6, and 10) binding reactions containing a 50-fold molar excess of unlabeled, nonspecific competitor DNA in the form of a double-stranded, synthetic oligonucleotide; (lanes 3 and 7, 11 and 12) binding reactions containing a 50-fold molar excess of unlabeled, specific oligonucleotides. The position of the protein–DNA complex α is indicated by an arrow.
nuclear extract or purified PTFl protein and synthetic oligonucleotide Amy2-IV, which encompasses the entire α-amylase gene cognate sequence, yield two complexes, α and β, in gel retardation assays. Although distinct by their electrophoretic mobility, the α and β complexes contain the same DNA-binding activity, as judged from footprint analysis or G methylation protection and interference studies (Cockell et al. 1989). The different electrophoretic behavior of these complexes most likely reflects a difference in protein mass rather than conformation because the activities generating α and β complexes can be separated by gel filtration on Sephacryl S-300 columns.

To elucidate whether A or B box sequences alone are sufficient to bind PTFl, we carried out gel retardation assays, using truncated versions of the PTFl cognate sequence and purified α or β forms of PTFl. Fractions were analyzed separately for their ability to bind to various oligonucleotide probes. The gel retardation analysis presented in Figure 1B shows the result with the protein fraction giving rise to α complex. Oligonucleotides representing truncated versions of the PTFl cognate sequence, Amy2-IVA (which lacks the B box) and Amy2-IVB (which lacks the A box), each form a single protein—DNA complex that has the same electrophoretic mobility as the one formed with Amy2-IV. However, the amount of complex generated by partial PTFl cognate sequences is reduced. The formation of the complex is nevertheless sequence specific, as it is competed preferentially by specific, rather than nonspecific, DNA sequence. This result is in line with earlier studies showing that oligonucleotides containing either A or B box sequences competed with PTFl footprint with efficiencies that were at least one order of magnitude lower than those obtained with the complete binding site (Cockell et al. 1989). The low efficiency of complex formation with truncated PTFl cognate sequences can be explained by their lower relative affinity for the protein. For instance, a 50-fold molar excess of the Amy2-IV oligonucleotide inhibits the formation of Amy2-IVB complex completely, whereas the same excess of Amy2-IVB oligonucleotide competes Amy2-IV complex only partially. Similar results were obtained with the protein of the fraction giving rise to complex β (data not shown). Again, the partial PTFl sequences yielded a single complex that had the same mobility as that formed with the entire PTFl recognition motif. The results from these experiments demonstrate that a tandem array of A and B box sequences is necessary for high-affinity binding of PTFl. The fact that DNA elements containing either box alone generate the protein–DNA complex expected for the full-length DNA sequence implies that PTFl binds as an entity to partial elements. This conclusion is supported by the observation that the Amy2-IVB DNA–protein complex is fully competed by the Amy2-IVB oligonucleotide (lane 12), which has no significant homology to the Amy2-IVA sequence. Two alternative possibilities concerning the structure of PTFl conform with these observations. One assumes the binding of a single polypeptide chain, the other one, a protein composed of subunits.

UV cross-linking identifies the DNA-binding components of PTFl

We performed UV cross-linking experiments to elucidate the structure of PTFl. The principle of this method is to label a protein selectively through its DNA-binding properties. The various DNA probes used for cross-linking were prepared by annealing a short primer to single strands of Amy2-IV DNA. Second strand synthesis then was carried out in the presence of [32P]dATP and azido-dUTP. DNA molecules substituted with the thymidine analog N3-dU cross-link with efficiencies that greatly exceed those achieved with templates substituted with bromodeoxyuridine (Evans and Haley 1987). The basic strategy used for cross-linking PTFl and Amy2-IV DNA is outlined in Figure 2. N3-dU-substituted, 32P-labeled DNA probes were incubated for binding with either crude nuclear proteins or with purified PTFl factor. In some experiments a proteolytic digestion of complexes was carried out prior to UV irradiation. The amount of specific protein–DNA complex formed and the extent of proteolysis were monitored by analyzing aliquots of the binding reactions in gel retardation assays. Protein–DNA complexes were cross-linked by exposure to short-wave UV light and analyzed by electrophoresis on denaturing sodium dodecyl sulfate (SDS) gels. Removal of DNA from complexes by micro-

Figure 2. The strategy for structural analysis of PTFl by site-directed UV cross-linking.
cocal nuclease preceded SDS-polyacrylamide gel electrophoresis (PAGE) analysis.

A protein that cross-links with Amy2-IV DNA has to fulfill a number of requirements to qualify as a DNA-binding component of PTFl. It must be sequence specific and the constituent of a cell-specific DNA-binding activity. Furthermore, we may expect it to be an integral part of both α and β protein–DNA complexes. The DNA probe used to address these issues was the Amy2-IV oligonucleotide, in which either the coding or noncoding strand had been fully substituted with N3·dU [Fig. 3A]. The cell specificity of protein–DNA cross-links was tested by carrying out binding reactions with crude nuclear extracts derived from two rat cell lines that originate from the same transplantable acinar pancreatic tumor. Cells from one of these lines (AR42J) express pancreas-specific functions, including PTFl DNA-binding activity, whereas cells from the other line (AR4IP) do not [Jessop and Hay 1980; Cockell et al. 1989]. The gel retardation assays of Figure 3B show that substitution of thymidine residues by N3·dU sustains the binding of PTFl to the probes and confirms that α and β complexes are cell specific. UV irradiation of the binding reactions containing extract of AR42J cells produces one predominant cross-linked protein–DNA complex with each DNA probe [Fig. 3B]. We designated the complexes generated with substituted coding and noncoding strands X and Y, respectively. X and Y, which differ in electrophoretic mobility, are not detected with nuclear extract from AR4IP cells [Fig. 3B].

The DNA sequence specificity of the proteins that yield X and Y complexes was established by comparing the results from binding reactions carried out with purified PTFl in the presence of an excess of unlabeled specific or nonspecific competitor DNA. The gel retardation assay of Figure 4 shows that formation of protein–DNA complexes α and β is specifically abolished by 50-fold molar excess of unlabeled Amy2-IV sequence. As

![Figure 3](genesdev.cshlp.org)
Transcription factor PTF1 is a heteromeric oligomer

identifies the proteins that generate X and Y complexes as DNA-binding components of PTF1.

PTF1 contains two different proteins that interact with the DNA

To determine whether the difference in mobility of X and Y complexes is due to the size of cross-linked protein, we have studied the electrophoretic behavior of these complexes, following digestion with micrococcal nuclease [Fig. 6]. We find that treatment by the enzyme affects the migration of both complexes but that the difference in mobility of X and Y persists even after extensive nuclease digestion. An apparent molecular mass of 64 and 48 kD can be estimated for X and Y, respectively, under conditions where most of the DNA is removed, as judged by a drastic reduction in intensity of labeled complexes.

Partial proteolytic digestion by sequence-specific endopeptidases can serve as a criterion to distinguish proteins that differ in amino acid sequence. We used Staphylococcus aureus V8 protease to compare the proteolytic patterns of X and Y polypeptides (Fig. 7). The experiment was done in the following way. Purified PTF1 was bound in separate reactions to Amy2-IV DNA

Figure 4. The DNA sequence specificity of cross-linked proteins. PTF1, purified by DNA affinity chromatography, was incubated in the presence or absence of specific or nonspecific competitor DNAs with the differentially substituted Amy2-IV probes of Fig. 3A. [Lanes 1 and 2] Standard binding reactions with no competitor added; [lanes 3 and 4] binding reactions containing a 50-fold molar excess of unlabeled Amy2-IV oligonucleotide; [lanes 5 and 6] binding reactions containing a 50-fold molar excess of a nonspecific oligonucleotide as competitor. Aliquots of the binding reactions and UV-irradiated samples were analyzed as described in Fig. 3B.

a consequence of this specific competition, cross-linked complexes X and Y are not formed. In contrast, competition with nonspecific DNA sequence does not affect the formation of either α and β or cross-linked complexes. The sequence-specific cross-linked complexes detected in binding reactions with purified PTF1 have the same relative electrophoretic mobility as those formed with crude nuclear extract.

The structural relationship between protein–DNA complexes detected by gel retardation analysis and by UV cross-linking was elucidated by carrying out cross-linking with the separated DNA-binding components of α and β complexes (Fig. 5). The α and β forms of PTF1 were separated and purified, as described for the experiment of Figure 1. The gel retardation assay shows that α and β fractions were not cross contaminated. Cross-linking of both α and β complexes results in the formation of X plus Y complexes.

The experimental evidence presented in Figures 3–5

Figure 5. Comparison of PTF1 proteins that generate α and β complexes. The α and β forms of PTF1 were separated by chromatography of crude nuclear extract from AR42J cells on a Sephacryl S-300 column. Fractions containing either the α [fraction I] or β [fraction II] form were incubated separately with the differentially substituted Amy2-IV probes of Fig. 3A. Aliquots of the binding reactions and UV-irradiated samples were analyzed as described in Fig. 3B.

GENES & DEVELOPMENT 1617
Figure 6. The effect of micrococcal nuclease on the electrophoretic mobility of cross-linked PTF1-DNA complexes. Binding reactions containing purified PTF1 and Amy2-IV probes of Fig. 3A were UV-irradiated. Complexes containing substituted coding (c) or noncoding (nc) Amy2-IV DNA strands were digested separately with two different concentrations of micrococcal nuclease, each prior to SDS-gel electrophoresis. The sizes of proteins generating complexes X and Y (arrows) were estimated from comparison to \(^{14}\)C-labeled molecular weight markers under conditions where most of the DNA was removed.

probes in which either the coding or the noncoding strand had been fully substituted with \(N_3\cdot dU\). One-half of each binding reaction was digested partially with the enzyme. The extent of digestion was monitored by gel retardation analysis and found to be very similar for complexes containing substituted coding or noncoding DNA strands. Then samples were UV-irradiated and separated on a denaturing SDS gel. We observe that the size of the 64-kD polypeptide, which generates X complex and cross-links with the coding strand, is not affected by the protease treatment. In contrast, the 48-kD polypeptide, which yields Y complex and cross-links with the noncoding strand, is sensitive to attack by the enzyme. This is evidenced by the increased electrophoretic mobility of the V8-resistant portion of Y (\(Y'\)). The fact that the two polypeptides can be distinguished by their size and sensitivity to V8 protease suggests that they differ in primary sequences.

The two polypeptides recognize different sequence motifs of the PTF1 cognate sequence and bind to the same DNA molecule

The bipartite arrangement of the DNA sequences that interact with PTF1 factor suggested to us that the 64- and 48-kD polypeptides recognize different domains of the PTF1 cognate sequence. Therefore, we designed an experiment with the aim of localizing sites that interact with the two polypeptides. For this purpose, we prepared a new set of Amy2-IV probes containing an A or B box-

specific substitution of thymidine by \(N_3\cdot dU\) (Fig. 8A). Gel analysis of the binding reactions containing purified PTF1 shows that the A box probes cross-link exclusively with the protease-resistant 64-kD protein, whereas the B box probe cross-links only with the protease-sensitive 48-kD protein (Fig. 8B). The cross-linking efficiency with \(c_A\) and \(nc_B\) probes is comparable to that obtained with the fully substituted probes, whereas the cross-linking efficiency with the \(nc_A\) probe is an order of magnitude lower (see legend to Fig. 8). Complexes that contain the 64-kD protein cross-linked to the A box-noncoding strand are also detected with a fully substituted probe (see, e.g., the weak band in lane nc of Fig. 3B).

The experiment of Figure 8 unambiguously assigns the two polypeptides to those DNA domains in which methylation of G residues interferes with PTF1 binding [Cockell et al. 1989]. However, the experiment, itself, does not provide evidence for the simultaneous presence of both polypeptides on a single DNA molecule. Therefore, we performed an experiment demonstrating directly that both polypeptides bind to the same DNA molecule. This necessitated synthesis of another Amy2-IV DNA probe in which the A and B box sequences were simultaneously substituted with \(N_3\cdot dU\), and both strands were labeled with \(^{32}\)P (Fig. 9A). Gel retardation

Figure 7. The effect of V8 protease on the electrophoretic mobility of cross-linked PTF1–DNA complexes. The differentially substituted Amy2-IV probes of Fig. 3A were incubated separately for binding of purified PTF1. Half of each binding reaction was digested with \(S. aureus\) V8 protease. Aliquots of digested and undigested samples were analyzed on a 2% agarose gel. The bulk of the samples was UV-irradiated and analyzed by SDS-PAGE. The positions of X and Y protein–DNA complexes are indicated by arrows. Digestion of the Y complex with the enzyme gives rise to the faster migrating form Y’ (open arrow).
Transcription factor PTF1 is a heteromeric oligomer.

**Figure 8.** Localization of the sites at which the two PTF1 proteins interact with Amy2-IV DNA. (A) Maps of the Amy2-IV DNA probes used to demonstrate the sites of interaction with PTF1. In the cA probe, thymidine residues 9, 10, and 11 (see Fig. 1A) of the coding strand have been substituted with N3-dU. In the ncA probe, thymidine residue 2 of the noncoding strand has been substituted. In the ncB probe, thymidines 9 and 10 of the noncoding strand have been replaced by N3-dU. For description of symbols, refer to Fig. 3A. (B) Binding reactions containing purified PTF1 and the probes depicted in A were subjected to digestion by V8 protease under conditions described in the legend to Fig. 7. Note that 10 times more material from the binding reaction containing the ncA DNA was used for SDS-PAGE analysis to compensate for the low cross-linking efficiency with this probe (for further explanation, see text).

Analysis of a binding reaction containing this probe cross-linked with purified PTF1 reveals two bands (upper and lower) that are resistant to denaturation by SDS (Fig. 9B, lane 3). The presence of these bands, which have an altered electrophoretic mobility as compared to α and β complexes (lane 1), is strictly dependent on UV-induced cross-links prior to SDS treatment. If the UV irradiation step is omitted, these bands are not generated, as α and β complexes are completely dissociated by the denaturing agent (lane 2). We expect only a minor fraction of the DNA molecules to be cross-linked simultaneously with the two polypeptides due to the low efficiency of cross-links generated. Therefore, the weak upper band in lane 3 is a candidate for this type of protein–DNA complex, whereas the strong lower band is expected to contain a mixture of DNA molecules cross-linked with either one or the other polypeptide. This prediction is confirmed by analysis of gel-purified upper and lower band complexes (Fig. 9C). Exposure of the upper band complex to high temperature, which results in strand separation of the DNA template, generates two complexes that comigrate with X and Y complexes derived from the lower band (cf. lanes 2 and 3). If the heat denaturation step is not carried out, the upper band complex migrates to a position that corresponds approximately to the sum of X and Y complexes [X + Y in lane 1] in molecular weight.

The results from this experiment and the one indicating that PTF1 binds as an entity to DNA templates containing the binding site for one of the two polypeptides only (see Fig. 1) suggest a subunit structure of PTF1. This interpretation of the data predicts that both the 64- and 48-kD polypeptides of PTF1 should be retained on affinity columns containing a truncated PTF1-binding site. Figure 10 illustrates that this is indeed the case. The experiment was performed in the following
Figure 9. Identification of a complex containing the two PTF1 proteins cross-linked with the same DNA molecule. (A) Map of the Amy2-IV DNA probe in which thymidines 9 and 10 of the noncoding strand and thymidines 9, 10, and 11 of the coding strand have been substituted with N3- dU (for synthesis, see Materials and methods; for description of symbols, refer to Fig. 3A). (B) Gel retardation analysis. The Amy2-IV DNA probe containing substitution of thymidine by N3- dU on both strands (A) was incubated with purified PTF1. The resulting protein–DNA complexes were subjected to SDS and/or UV light and were then analyzed on a 1.8% low-melt agarose gel. (Lane 1) Standard binding reaction as a control; (lane 2) the binding reaction was treated with 0.1% SDS prior to electrophoresis; (lane 3) the binding reaction was UV-irradiated prior to treatment with 0.1% SDS. Two bands (upper and lower) representing cross-linked protein–DNA complexes are detected in lane 3. (C) SDS-PAGE analysis. The upper and lower band complexes of lane 3 (B) were eluted from the gel by diffusion overnight at room temperature in SDS-PAGE loading buffer and concentrated. Half of the sample containing upper band complex was loaded directly onto the gel without any further treatment. The other half and the sample containing the lower band complex were heated prior to electrophoresis for 2 min at 90°C for strand separation of the DNA template. The higher labeling intensity of complex X is due to a threefold higher specific activity of the coding DNA strand.

way. Crude nuclear extract of AR42J cells was subjected to chromatography on a column bearing the Amy2-IV oligonucleotide, which contains only the recognition motif for the 48-kD protein. PTF1 is expected to bind inefficiently to this column because of its low affinity for truncated binding sites [see Fig. 1]. Indeed, gel retardation analysis shows that most of the PTF1 DNA-binding activity is recovered from the flowthrough fraction of the column [Fig. 10A, lane 2]. A binding reaction with column-bound material reveals formation of a protein–DNA complex having the electrophoretic mobility of PTF1-specific complex α [Fig. 10A, lane 4]. In addition, we observe a novel complex, δ, which migrates slightly faster in the gel than PTF1-specific complex β. The proteins generating these two complexes were identified after UV cross-linking to 32P-labeled Amy2-IV oligonucleotides, in which either the A or the B box had been substituted with N3- dU. Complexes α and δ from these binding reactions were gel-purified and analyzed separately on a denaturing polyacrylamide gel. Figure 10B shows that cross-linking of complex α yields X plus Y complexes, which are diagnostic for the 64- and 48-kD polypeptides of PTF1, respectively [cf. lanes 1 and 2]. From this we conclude that PTF1 binds as an entity to an affinity column containing the binding site for the 48-kD protein only. Cross-linking of complex δ yields complex Z, which is distinct by its electrophoretic behavior from PTF1-specific complex Y. We take this observation to indicate that Z and Y complexes are generated by different proteins and that free forms of the 48-kD protein, if they exist, do not bind to the column.

The combined evidence from experiments shown in
Transcription factor PTF1 is a heteromeric oligomer

Discussion

We analyzed the structure of the pancreas-specific transcription factor PTF1 by site-directed UV cross-linking. The results from our studies are summarized in Figure 11, which shows that PTF1 interacts with DNA through two distinct proteins. These proteins differ in all the properties we examined. They have different apparent molecular weights, different sensitivity to V8 protease, and recognize different DNA sequence motifs. Apparently, the two proteins act in concert to confer DNA-binding activity to PTF1 because we do not observe complexes containing either protein alone bound to DNA. However, a complex containing both proteins may bind, although with a markedly reduced efficiency, to DNA containing the recognition site for only one of the proteins. These observations suggest that dimerization of the two proteins is required for DNA binding but do not rule out the possibility that free forms of these proteins having little or no affinity for PTF1 recognition motifs may exist also. Independent evidence, which shows that both proteins bind to the same molecule of DNA, supports the notion that PTF1 binds as a single entity to DNA. Our data are compatible with a model where PTF1 is an oligomer containing two different subunits that both interact with the DNA. We prefer to use the term heteromeric oligomer rather than heterodimer to describe the structure of PTF1 because gel electrophoresis and gel filtration chromatography both suggest that PTF1 complexes α and β may differ in protein content.

The subunit structure of PTF1 is reflected at the DNA level in a bipartite organization of the PTF1-binding site that comprises two distinct recognition motifs (Cockell et al. 1989). The centers of the two domains of protein–DNA interaction are separated by two helical turns of DNA in the α-amylase gene element, whereas they are spaced by one helical turn of DNA in other genes expressed in the acinar pancreas. The two domains of PTF1-binding sites not only vary in their spatial distribution but also exhibit sequence variability among different genes. UV cross-linking experiments performed with binding sites of trypsin 1 and elastase 2 genes and PTF1, which was purified on affinity columns bearing the α-amylase gene-binding site, disclose two proteins having properties indistinguishable from those of proteins binding to the α-amylase gene element (data not shown). Therefore, the arguments made in this paper with respect to the interaction of PTF1 with the α-amylase cognate sequence can be extended to include cognate sequences of these other genes as well. These observations suggest that PTF1 is able to recognize DNA elements that differ in the respective nucleotide sequences of the two domains of protein–DNA interaction which, moreover, can have different spatial arrangements.

PTF1 is distinct from other multimeric transcription...
Figure 11. The structure of the PTFl-DNA complex. The model is based primarily on results from site-directed UV cross-linking experiments of this paper but it also incorporates data from other studies (Cockell et al. 1989). (●) The putative positions at which substitution of T residues by N3-dU generates protein-DNA cross-links. (A) A box; (B) B box; (nc) noncoding DNA strand; (c) coding DNA strand.

Factors in that it contains two heterologous DNA-binding subunits. Several transcription factors have been reported to exist as multimers, either in solution or when complexed with DNA. Representative examples of homodimeric factors are the GCN4 protein of yeast, which dimerizes in solution (Hope and Struhl 1986), and receptors for glucocorticoid and progesterone hormones, which dimerize upon binding to the DNA (Tsai et al. 1988). Two examples of heterodimeric transcription factors in which only one of the subunits exhibits DNA-binding activity have been described. Both factors belong to a class of proteins that recognize CCAAT motifs in promoters of genes from yeast and man. The yeast factor, which activates transcription from the iso-1 cytochrome c gene promoter, has been shown, by genetic and biochemical analysis, to exist as a complex containing the gene products of the HAP2 and HAP3 loci (Olesen et al. 1987; Hahn and Guarente 1988). A similar situation applies to human CCAAT-binding factor CPI. The human factor is apparently closely related to that of yeast because its two subunits, CPI A and CPI B, functionally substitute for the respective subunits of yeast factor in binding assays in vitro (Chodosh et al. 1988). Recently, another type of heterodimeric transcription factor has been described. Dimerization of c-jun and c-fos oncoproteins activates transcription from gene promoters containing AP-1-binding sites (Halazonetis et al. 1987; Nakabeppu et al. 1988). The heterodimer has a greatly increased DNA-binding affinity as compared to the c-jun homodimer.

We can only speculate as to the reasons why the peculiar heteromeric structure of PTFl has arisen during evolution. We may consider the interaction of PTFl with DNA through two subunits that bind in an independent manner as an extreme form of cooperative binding. Originally, the two subunits may have existed as two separate DNA-binding proteins, the interplay of which was essential for the function of promoters in a cell-specific manner. Such combinatorial mechanisms have been implicated in the expression of genes encoding liver-specific functions (Cereghini et al. 1987; Lichtsteiner et al. 1987; Hardon et al. 1988). In light of the apparent importance of such mechanisms for cell-specific gene expression, we may postulate that the structure of PTFl reflects a combinatorial event that has become permanently fixed during evolution. Binding sites for PTFl exist in front of many genes expressed in the acinar pancreas as part of DNA sequences that have been defined as transcriptional enhancers (Boulet et al. 1986; Kruse et al. 1988; Cockell et al. 1989). From this observation, we may extrapolate that PTFl is part of the mechanism responsible for the coordinate expression of these genes. In this context, it will be of interest to determine the cellular distribution and the developmental profiles of the mRNAs encoding PTFl proteins.

Materials and methods

Tissue culture conditions and preparation of nuclear extracts

AR42J and AR4IP rat cells (Jessop and Hay 1980) were cultured, as described by Cockell et al. (1989). Crude nuclear extracts were prepared from these cells, as described by Dignam et al. (1983), except that cell homogenization and nuclear lysis buffers also contained a mixture of specific and nonspecific protease inhibitors (Cockell et al. 1989).

Chromatographic procedures

Separation of α and β forms of PTFl was accomplished by chromatography on Sephacryl S-300 columns (Pharmacia). One milligram of nuclear extract from AR42J cells was loaded on a 50 × 0.7-cm column that had been equilibrated in 12% glycerol, 12 mm HEPES [pH 7.9], 0.4 M KCl, 0.12 mm EDTA, 5 mm dithiothreitol [DTT], and 0.5% Triton X-100 prior to use. The sample was loaded in the same buffer containing 20% glycerol.

Crude nuclear extract was subjected to two cycles of DNA affinity column chromatography, as described by Cockell et al. (1989). The Amy2-IV oligonucleotide, which contains the PTFl cognate sequence of the α-amylase gene, was biotinylated and immobilized on a streptavidin-agarose matrix, according to the procedure described by Chodosh et al. (1986). This proce...
Preparation of DNA probes for UV cross-linking

Oligonucleotides were synthesized on an Applied Biosystems DNA synthesizer and purified on 15% sequencing gels prior to use. Double-stranded DNA probes were formed by annealing complementary oligonucleotides at a concentration of 50 µg/ml each in 10 mM Tris-HCl (pH 7.9) and 5 mM MgCl₂. The annealing mixture was heated for 5 min at 70°C and then cooled slowly to room temperature.

Six different DNA probes containing localized substitution of thymidine by N₃·dU (donated to us by R.K. Evans) were used for cross-linking experiments described in this paper. The synthesis and properties of N₃·dUTP have been reported in detail by Evans and Haley (1987).

1. Synthesis of N₃·dU/c and N₃·dU/nc (maps shown in Fig. 3A) The double-stranded Amy2-IV oligonucleotide (see Fig. 1A) was cloned into the BamHI site of SP65 vector DNA. A DNA fragment containing Amy2-IV plus polylinker sequences was excised from the vector by cleavage with EcoRI and HindIII. This cleavage, which generates an asymmetry in size of the two DNA strands, allowed strand separation of the fragment on an 8% sequencing gel. The isolated DNA strands then were annealed to strand-specific primers complementary to polylinker sequences at their 3' ends. Double-stranded probes were prepared by primer-extended second strand synthesis, using Klenow enzyme in the presence of 100 µM dGTP, dCTP, N₃·dUTP, and 10 µM [³²P]dATP (3000 Ci/m mole). Then the DNA was digested with BamHI, and polylinker sequences were removed by electrophoresis on a 6% polyacrylamide gel.

2. Synthesis of N₃·dU/cA (map of Fig. 8A) The noncoding strand of synthetic Amy2-IV oligonucleotide was annealed to a primer complementary to sequences -120 to -141 [see Fig. 1A]. Coding strand synthesis in the presence of N₃·dUTP and [³²P]dATP was as described in 1 (above).

3. Synthesis of N₃·dU/ncA (map of Fig. 8A) The coding strand used for synthesis of the N₃·dU/nc probe [see 1, above] was annealed to a primer complementary to sequences of the polylinker. Noncoding strand synthesis was initiated at nucleotide position -154 by Klenow enzyme and was carried out in two steps. The first step, in the presence of [³²P]dATP and N₃·dUTP, allows the addition of 2 nucleotides to the primer and results in the substitution of T residue 2 [see Fig. 1A]. The dNTPs were then removed, [³²P]dATP, dGTP, dCTP, and dTTP were freshly added, and noncoding strand synthesis was completed. Then the DNA was digested with BamHI, and polylinker sequences were removed by electrophoresis on a polyacrylamide gel.

4. Synthesis of N₃·dU/ncB (map of Fig. 8A) The synthetic oligonucleotide bearing coding strand sequences of Amy2-IV DNA was modified first at its 3' end by addition of a dideoxynucleotide using terminal transferase to prevent elongation from this end during second strand synthesis. The modified strand then was annealed to a primer complementary to sequences -130 and -156 [see Fig. 1A]. Noncoding strand synthesis by Klenow enzyme then was carried out in two steps. Synthesis was initiated in the presence of dGTP, dCTP, and N₃·dUTP. This step allows the addition of 6 nucleotides to the primer and results in the incorporation of N₃·dU at positions corre-
sponding to T residues 9 and 10 [see Fig. 1A]. Then the dNTPs were removed, [³²P]dATP, dGTP, dCTP, and dTTP were freshly added, and noncoding strand synthesis was completed.

5. Synthesis of N₃·dU/cA + ncA (map of Fig. 9) An oligonucleotide containing coding strand sequences -120 to -141 of Amy2-IV was annealed to an oligonucleotide bearing noncoding strand sequences -130 to -156 of Amy2-IV. The recessed 3' ends of this DNA molecule then were filled in as described in 1 (above).

Gel retardation assays

Binding reactions [20 µl] containing 0.05 ng of [³²P]-labeled DNA and 10 µg of crude nuclear extract were incubated in 12% glycerol, 12 mM HEPES (pH 7.9), 60 mM KCl, 0.12 mM EDTA, and 5 mM MgCl₂ [binding buffer] for 20 min at room temperature. DTT was omitted from the binding buffer because it inactivates the cross-linking agent [Evans and Haley 1987]. All binding reactions containing crude nuclear extract were carried out in the presence of 2 µg of single-stranded E. coli DNA as nonspecific competitor. Binding reactions with DNA affinity column fractions contained 2.5 ng of nonspecific DNA competitor plus 250 µg/ml of BSA. Competition experiments with synthetic, unlabeled oligonucleotides were done by simultaneously adding a 50-fold molar excess of competing sequence and the radioactive probe. Proteolytic digestion of the binding reactions with S. aureus V8 protease was carried out for 10 min at room temperature in binding buffer and at an enzyme concentration of 7.5 µg/ml. Binding reactions were subjected to electrophoresis for 2 hr at 40 mA and 0°C on 2% agarose gels in 0.5 × TBE buffer (pH 8.3). Protein–DNA complexes then were detected by autoradiography.

UV cross-linking

The binding reactions (in Eppendorf tubes) were irradiated for 10–30 min on ice at a distance of 10 cm, using a UV transilluminator (254 nm, 7000 µW/cm²). Under these conditions, ~5–15% of total protein–DNA complex is cross-linked. Digestion with micrococcal nuclease (Worthington Diagnostics, Freehold, New Jersey) was carried out at 37°C for 10 min in the presence of 10 mM CaCl₂. No contaminating protease activities were detected in the particular batch of enzyme used. Cross-linked material was separated on 10% SDS–polyacrylamide gels (Laemmli 1970) and visualized by autoradiography.

Acknowledgments

We are greatly indebted to Dr. R.K. Evans for a generous gift of azido-dUTP. We thank Dr. L. Kühn for oligonucleotide synthesis and R. Bovey for expert technical assistance. We also thank M. Cockell for critical reading of the manuscript. O.H. was supported by a career developmental award of the Cloëtta Foundation. This work was supported by a grant from the Swiss National Science Foundation.

References

Boulou, 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.

Cereghini, S., M. Raymondjean, A.G. Carranca, P. Herbomel, and M. Yaniv. 1987. Factors involved in control of tissue-specific expression of albumin gene. Cell 50: 627–638.
Roux et al.

Chodosh, L.A., R.W. Carthew, and P.A. Sharp. 1986. A single polypeptide possesses the binding and transcription activities of the adenovirus major late transcription factor. *Mol. Cell. Biol.* 6: 4723–4733.

Chodosh, L.A., J. Olesen, S. Hahn, A.S. Baldwin, L. Guarente, and P.A. Sharp. 1988. A yeast and a human CCAAT-binding protein have heterologous subunits that are functionally interchangeable. *Cell* 53: 25–35.

Cockell, M., B.J. Stevenson, M. Strubin, O. Hagenbüchle, 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.

Dignam, J.D., R.M. Lebovitz, and R.G. Roeder. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* 11: 1475–1489.

Evans, R.K. and B.E. Haley. 1987. Synthesis and biological properties of 5-azido-2'-deoxyuridine 5'-triphosphate, a photoactive nucleotide suitable for making light-sensitive DNA. *Biochemistry* 26: 269–276.

Hahn, S. and L. Guarente. 1988. Yeast HAP2 and HAP3: Transcriptional activators in a heteromeric complex. *Science* 240: 317–321.

Halazonetis, T.D., K. Georgopoulos, M.E. Greenberg, and P. Leder. 1988. C-Jun dimerizes with itself and with c-Fos, forming complexes of different DNA binding affinities. *Cell* 55: 917–924.

Hardon, E.M., M. Frain, G. Paonessa, and R. Cortese. 1988. Two distinct factors interact with the promoter regions of several liver specific genes. *EMBO J.* 7: 1711-1719.

Hope, I.A. and K. Struhl. 1986. Functional dissection of a eukaryotic transcription activator protein, GCN4 of yeast. *Cell* 46: 885–894.

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

Kruse, F., C. Komro, C. Michnöf, and R. MacDonald. 1988. The cell-specific elastase I enhancer comprises two domains. *Mol. Cell. Biol.* 13: 893–902.

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680–685.

Lichtsteiner, S., J. Wuarin, and U. Schibbler. 1987. The interplay of DNA-binding proteins on the promoter of the mouse albumin gene. *Cell* 51: 963–973.

Nakabeppu, Y., K. Ryder, and D. Nathans. 1988. DNA binding activities of three murine jun proteins: Stimulation by fos. *Cell* 55: 907–915.

Olesen, J., S. Hahn, and L. Guarente. 1987. Yeast HAP2 and HAP3 activators both bind to the CYC1 upstream activation site, UAS2, in an interdependent manner. *Cell* 51: 953–961.

Tsai, S.Y., J. Carlstedt-Duke, N.L. Weigel, K. Dahlmann, J.-A. Gustafsson, M.H. Tsai, and B.W. O'Malley. 1988. Molecular interactions of steroid hormone receptor with its enhancer element: Evidence for receptor dimer formation. *Cell* 55: 361–369.
The cell-specific transcription factor PTF1 contains two different subunits that interact with the DNA.

E Roux, M Strubin, O Hagenbüchle, et al.

*Genes Dev.* 1989, 3:
Access the most recent version at doi:10.1101/gad.3.10.1613