Structural and functional properties of the N transcriptional activation domain of thyroid transcription factor-1: similarities with the acidic activation domains

Gianluca TELL*[1], Lorena PERRONE†, Dora FABBRO*, Lucia PELLIZZARI*, Carlo PUCILLO*, Mario DE FELICE†‡, Renato ACQUAVIVA§, Silvestro FORMISANO§ and Giuseppe DAMANTE*

*Dipartimento di Scienze e Tecnologie Biomediche, Università degli Studi di Udine, Via Gervasutta 48, 33100 Udine, Italy, †Stazione Zoologica 'A. Dohrn', Villa Comunale 1, 80121 Napoli, Italy, ‡Istituto Pluridisciplinare di Patologia Generale, Facoltà di Medicina, Università di Messina, 98100 Messina, Italy, and §C.E.O.S., c/o Dipartimento di Biologia e Patologia Cellulare e Molecolare 'L. Califano', Università degli Studi di Napoli 'Federico II', via Pansini 5, 80131 Napoli, Italy

The thyroid transcription factor 1 (TTF-1) is a tissue-specific transcription factor involved in the development of thyroid and lung. TTF-1 contains two transcriptional activation domains (N and C domain). The primary amino acid sequence of the N domain does not show any typical characteristic of known transcriptional activation domains. In aqueous solution the N domain exists in a random-coil conformation. The increase of the milieu hydrophobicity, by the addition of trifluoroethanol, induces a considerable gain of α-helical structure. Acidic transcriptional activation domains are largely unstructured in solution, but, under hydrophobic conditions, folding into α-helices or β-strands can be induced. Therefore our data indicate that the inducibility of α-helix by hydrophobic conditions is a property not restricted to acidic domains. Co-transfections experiments indicate that the acidic domain of herpes simplex virus protein VP16 (VP16) and the TTF-1 N domain are interchangeable and that a chimaeric protein, which combines VP16 linked to the DNA-binding domain of TTF-1, undergoes the same regulatory constraints that operate for the wild-type TTF-1. In addition, we demonstrate that the TTF-1 N domain possesses two typical properties of acidic activation domains: TBP (TATA-binding protein) binding and ability to activate transcription in yeast. Accordingly, the TTF-1 N domain is able to squelch the activity of the p65 acidic domain. Altogether, these structural and functional data suggest that a non-acidic transcriptional activation domain (TTF-1 N domain) activates transcription by using molecular mechanisms similar to those used by acidic domains. TTF-1 N domain and acidic domains define a family of proteins whose common property is to activate transcription through the use of mechanisms largely conserved during evolutionary development.

INTRODUCTION

Promoter-specific transcription factors play an essential role in the regulation of gene expression [1]. These molecules are multidomain proteins. In fact, the two basic molecular functions of promoter-specific transcription factors, i.e. specific DNA binding and transcriptional activation, are performed by distinct, independent domains [2]. DNA-binding domains are very well known, and a large number of structural and functional studies reveal that these domains can be classified according to their three-dimensional structure [3]. Members of the same class display similar structural features, as well as a conserved procedure to recognize specific DNA sequences. Transcriptional activation domains are grouped according to the content of certain amino acids. Proline-rich, glutamine-rich and acidic domains are the most common ([2,4] and references cited therein). The presence of hydrophobic residues appears to be an essential feature for the function of these domains [5]. In contrast with DNA-binding domains (DBDs), information on the structure of transcriptional activation domains remains scanty and no structural models have been proposed to date which account for the role of these domains in the interactions with molecules of the transcriptional machinery. The acidic transcriptional-activation domains appear to be largely unstructured in aqueous solution, although under certain conditions some can adopt secondary structure [6–8]. The hydrophobic milieu, together with particular pH conditions, appear to be effective in inducing such secondary structure. It has been proposed that the ability of acidic activation domains to adopt a definite secondary structure (α-helix or β-sheet) may affect the activating function, via an 'induced fit' mechanism which drives the unstructured-to-structured transition upon contact with their target molecules [7].

Thyroid transcription factor 1 (TTF-1) is a tissue-specific protein which controls the transcription of some thyroid- and lung-specific genes [9]. However, very little is known about the mechanism by which TTF-1 activates the transcription of the thyroglobulin (Tg) and thyroperoxidase (TPO) genes only in follicular thyroid cells [10] and the transcription of the surfactant protein B (SPB) gene only in epithelial lung cells [11]. TTF-1 exhibits two independent transcriptional activation domains, located at the N- (N domain) and C-terminal (C domain) regions with respect to the DNA-contacting homoeodomain (HD) [12]. Although the N and C domains appear to be partially redundant, a functional hierarchy occurs between these domains. In fact, competition experiments have revealed that the N domain, but not the C domain, when present in a form unable to bind the target promoter, is able to squelch both its own transcriptional activity and that of the C domain [12]. These data would indicate

---

Abbreviations used: TTF-1, thyroid transcription factor-1; TBP, TATA-binding protein; CAT, chloramphenicol acetyltransferase; Tg, thyroglobulin; TPO, thyroperoxidase; SPB, surfactant protein B; HD, homoeodomain; DBD, DNA-binding domain; TFA, trifluoroacetic acid; TFE, trifluoroethanol; Ni-NTA, nickel/nitrilotriacetic acid; LB, Luria–Bertani; IPTG, isopropyl β-D-thiogalactoside; VP16, herpes simplex virus protein VP16; HDVP16, a chimaeric molecule containing the DNA-binding domain of TTF-1 linked to the acidic activation domain of VP16.

[1] To whom correspondence should be addressed.

that the N domain interacts with factor(s) essential for the activation of the basal transcriptional machinery. Contrary to what might be expected from its leading role, the primary sequence of the N domain (Figure 1 below, left-hand panel) does not contain typical characteristics of transcription-activation domains known up to now [6–8]. Thus the characterization of this domain could contribute to a better understanding of the principles of transcriptional activation.

In the present study, by means of structural and functional approaches, some features of the TTF-1 N domain have been delineated. The results suggest that, in activating transcription, this protein may use molecular mechanisms similar to those used by acidic activation domains.

MATERIALS AND METHODS

Bacterial expression and purification of the N domain

The cDNA coding for the amino acids 1–156 of rat TTF-1 was cloned into the vector pQ12 (Qiagen) in-frame with the coding region of six histidine residues. Thus the expressed protein (N domain–His$_6$) contains an extra hexahistidine sequence at the C-terminus, allowing for purification of the protein by nickel/nitrilotriacetic acid (Ni-NTA) affinity chromatography. The N domain–His$_6$ fusion protein, consisting of 162 residues plus an initial methionine residue derived from the starting codon, was expressed in M15 *Escherichia coli* cells.

Overnight cultures were inoculated into Luria-Bertani (LB) medium supplemented with 50 μg/ml ampicillin and grown to an A_{600} of 0.8 at 37 °C. Induction was achieved by adding isopropyl β-D-thiogalactoside (IPTG) to a concentration of 0.5 mM and incubating at 37 °C for an additional 3 h. Cells were harvested by centrifugation at 600 *g* in a Sorvall GS-3 rotor. The bacterial pellet was resuspended in lysis buffer A (6 M guanidinium chloride/0.1 M NaH$_2$PO$_4$/0.01 M Tris, pH 8.0) and centrifuged at 10000 *g* for 20 min at 10 °C. The supernatant was loaded on to an Ni-NTA column, equilibrated with buffer A, and washed with 10 vol. of buffer B (8 M urea/0.1 M NaH$_2$PO$_4$/0.01 M Tris, pH 8.0). The protein was eluted with buffer B (pH adjusted to 4.5). To prevent dimerization, the protein was acetylated at Cys^{87} and then dialysed against 10 mM sodium acetate buffer, pH 4.4. Sample concentrations were determined either spectrophotometrically (using ϵ_{278} 14000 M$^{-1}\cdot$cm^{-1} calculated as described by Wetlaufer [13]) or colorimetrically using Lowry et al. [14] or Bradford [15] assays. Absorbance measurements were obtained using a Perkin–Elmer Lambda 14 UV–visible spectrophotometer.

Expression and partial purification of recombinant TATA-binding protein (TBP)

TBP was purified using the procedure described by Brenowitz et al. [16] with minor modifications. *E. coli* BL21 (DE3) cells carrying the plasmids pLysS and pKA9, expressing yeast recombinant TBP, were grown at 37 °C in LB medium containing 30 μg/ml chloramphenicol and 25 μg/ml ampicillin. TBP expression was induced by the addition of 1mM IPTG when the cell density reached an A_{600} of 0.4. Induction was then performed at 30 °C for an additional 2 h. Cells were harvested by centrifugation and resuspended in lysis buffer [20 mM Hepes (pH 7.9)/1mM EDTA/1 mM EGTA/10 mM 2-mercapto-ethanol/2 μg/ml leupeptin/2 μg/ml pepstatin/1 mM PMSF] in a volume of 10 ml/g of bacterial pellet. After lysis by sonication, the cell debris was removed by centrifugation. DNA was removed by addition of protamine sulphate to the supernatant to 0.3 mg/ml and the precipitate removed by centrifugation. The supernatant was then dialysed against buffer D [20 mM Hepes

(pH 7.9)/20 % glycerol/1 mM EDTA] plus 100 mM KCl and loaded on to a Pharmacia S-Sepharose Fast Flow column pre-equilibrated with buffer D plus 100 mM KCl, washed with the same buffer and then eluted with a linear gradient from 100 mM to 700 mM KCl in buffer D. Fractions containing partially purified (10 %) TBP were stored at −70 °C.

CD

Jacketted cells of 0.2–0.5 cm pathlength were used, and typically ten spectra were accumulated, averaged and baseline-corrected on a Jasco J-600 spectropolarimeter interfaced to an Olidata personal computer. Calibration of the instrument was performed with D-(+)-10-camphorsulphonic acid at 290 nm. All spectra were collected at 25 $\pm$ 0.1 °C. The temperature was controlled by a Haake F3 water bath. A mean residue molecular mass of 103.2 Da was used for conversion into mean residue ellipticity ([θ], in degrees$\cdot$cm$^2\cdot$dmol^{-1}). The secondary-structure contents were calculated as described by Menendéz-Arias et al. [17]. pI was calculated by using the module CHARGPRO of the University of Wisconsin Genetics Computer Group software package.

Proteolysis with thermolysin

Horse heart cytochrome *c* (type IV) and thermolysin were obtained from Sigma. Trifluoroethanol (TFE) was purchased from Fluka, and the materials used for SDS/PAGE were from Sigma.

SDS/PAGE was carried out in a vertical-slab-gel apparatus (C.B.S. Scientific Co., Del Mar, CA, U.S.A.) using the poly-acrylamide system of Shagger and von Jagow [18]. Band detection was performed with Coomassie Brilliant Blue R-250. Digestion of proteins was performed by incubation of the protein dissolved (1 mg/ml) in 20 mM Tris/HCl, pH 7.8, containing 10 mM CaCl$_2$, with thermolysin (protease/protein substrate ratio 1:50, w/w) at 25 °C. The reaction was carried out in the absence or in the presence (10 or 50 %v/v) of TFE. The extent of proteolysis was assayed by SDS/PAGE analysis on 15 μl aliquots that were removed from the reaction mixture, added to 5 μl of 4 × Laemmli sample buffer and boiled for 5 min to inactivate the protease.

Cell cultures and transfection

HeLa cells were grown in Dulbecco's modified Eagle's medium supplemented with 10 % fetal-calf serum. For transient expression assays, cells were plated at 0.8 × 10^6 cells/100 mm tissue-culture dish 4–6 h before transfection. Transfections were carried out by the calcium phosphate co-precipitation technique [19]. The cells, after exposure to the calcium phosphate/DNA precipitate for 12–14 h, were washed with PBS and grown for 48 h. Chloramphenicol acetyltransferase (CAT) [20] and luciferase activities [21] in cells extracts were determined as described in the cited references. Plasmid pTACAT3 contains the wild-type Tg promoter linked to the CAT gene and is described in [22]. Mutants of Tg promoter A-core and C-core have been described in [23]. Plasmids with the promoters C5E1b and E1b, as well as plasmids expressing proteins ΔG21, Δ14 and Δ3, were described in [12]. The plasmid expressing the chimaeric protein HDVP16 was generated first by fusing the sequences coding for TTF-1 HD and herpes simplex virus protein VP16 (VP16) using the procedure of Ho et al. [24]. Subsequently the chimaeric gene has been inserted in the expressing vector Rc-CMV (Invitrogen Corp., San Diego, CA, U.S.A.). The p65-GAL4 expression vector was provided by Dr. P. De Luca (D.I.B.I.T., H. S. Raffaele, Milan, Italy). It is a derivative of pSG424 and contains the coding

sequence of the p65 activation found in-frame with the DBD of GAL4. The protein levels achieved in transfection experiments were monitored by gel-retardation assays [12].

Yeast strains and plasmids

Yeast strains used were HF7c [$MAT\alpha$, ura3-52, his3-200, ade2-101, lys2-801, trp1-901, leu2-3, 112, gal4-542, gal80-538, LYS2::GAL1-HIS3, URA3::(GAL4   17-mers)$_3$-CYC1-lacZ] (Clontech, Palo Alto, CA, U.S.A.), and INVSC1 ($MAT\alpha$, his3-$\Delta 1$, leu2, trp1-289, ura3-52) (Invitrogen). Yeast were grown in YEPD (yeast extract peptone dextrose) or selective minimal medium [25]. Transformations were made by the method of Schiestl and Gietz [26]. β-Galactosidase activity was assayed in liquid as described in [27]. TTF-1 deletion mutants were generated by PCR using primers encoding the amino acids at the ends of each fragment. The primers included either EcoRI or BamHI restriction sites to facilitate cloning. The deletions were cloned into the EcoRI and BamHI sites of the pGBT9 plasmid (Clontech) in order to obtain GAL4$_{(1-147)}$–TTF-1 fusion proteins. The Tg promoter fragment, from -170 to -23, was generated by PCR using primers including either SalI or XhoI. It was cloned in plasmid sx178 (obtained from Dr. G. Pullitzer, Department of Genetics, University of Naples, Naples, Italy), which is an upstream activating sequence lacking the CYC1::lacZ fusion plasmid derivative of pLG178 [27].

In vitro protein-binding assay

Precipitation experiments were performed using the TTF-1 N domain and BSA linked to CNBr-activated Sepharose 4B resin (Pharmacia) at a final concentration of 500 μg/ml of resin. CNBr-activated Sepharose 4B beads (100 μl) containing immobilized proteins were preincubated with binding buffer [20 mM Hepes (pH 7.9)/100 mM KCl/0.2 mM EDTA/0.2 mM EGTA/2 mM DTT/0.5 mM PMSF/20 % glycerol] for 2 h at 4 °C, then incubated with the load partially purified TBP fraction (250 μl) at 4 °C overnight and washed with 5 ml of binding buffer. Elution was carried out with 50 μl of 1M KCl . The eluted fractions were run either on a native gel (electrophoretic mobility-shift assay) as previously reported [12] using the sequence CYC1 as probe [28], and on a denaturing SDS/12 %-PAGE reducing gel. Polypeptides fractionated by SDS/PAGE were transferred to a nitrocellulose membrane (MSI, Westboro, MA, U.S.A.). The membrane was blocked in a mixture containing 5 % milk powder in PBS/0.1 % Tween-20. The TBP antibody was purchased from Santa Cruz Biotechnology, Santa Cruz, CA, U.S.A. The secondary antibody was horseradish peroxidase (HRP)-conjugated anti-rabbit IgG (Sigma). The membrane was developed with an Amersham ECL (enhanced chemiluminescence) detection kit and exposed to Amersham MP films.

RESULTS AND DISCUSSION

General properties and purification of TTF-1 N domain

The primary sequence of the N domain, along with predictions of its secondary structure, obtained by different methods [29], are reported in Figure 1 (left-hand panel) (see also Table 1). Statistical-information-based methods, like those described by Chou and Fasman [30], Garnier et al. [31] and Nagano and Hashegawa [32], led to similar results, with three or four helical tracts among the first 80 residues. The more accurate algorithm PHD developed by Rost and Sander [33], substantially agrees with these predictions, though the helical tracts are somewhat longer or even joined. The output of the algorithm adds further confidence in the tendency towards helix formation of this

region, by providing a reliability index which is for several residues top ranking. All methods predict a high content of α-helix ($> 75 \%$) in the residue 49–73 region. Interestingly the functional analysis of the N domain revealed that a critical area for the trans-activating function is found between residues 51 and 102 [12], which suggests that the predicted α-helical folding may play a role in activating the basal transcriptional machinery.

In order to investigate some structural features of TTF-1 N domain, the protein was expressed in bacteria as a His$_6$ fusion protein. The final yield of the protein was 30–60 mg from 1 litre of liquid culture. Figure 1 (middle panel) shows the HPLC chromatogram indicating that the achieved purity was over 98 %. The electrophoretic mobility (Figure 1, right-hand panel) implied an apparent molecular mass of 16.9 ± 0.7 kDa. This is in agreement with the expected value (16.8 kDa). The N domain–His$_6$ fusion protein proved sensitive to the redox potential of the environment. In fact, the non-acetylated protein was monomeric in the presence of 1 mM 2-mercaptoethanol (Figure 1, right-hand panel, lane 3) or 2 mM dithiothreitol (results not shown). In the absence of reducing agents, the protein could also be found in a dimeric form (Figure 1, right-hand panel, lane 4). It has been demonstrated that the in vitro DNA-binding activity of TTF-1 is regulated by the redox potential through oxidation of cysteine residues. Mutagenesis experiments revealed that at least two cysteine residues (at positions 87 and 363) are required for a redox-dependent modulation of the DNA-binding activity [34]. Our results, however, indicate that Cys[87] alone is sufficient for dimerization under the experimental conditions used in the present study. The functional relevance of a possible dimerization of TTF-1 involving only Cys[87] is not, as yet, understood.

CD spectroscopy

In contrast with the secondary-structure prediction analysis (Figure 1, left-hand panel), the CD spectrum of the purified N domain shows no apparent α-helical structure when analysed in aqueous solution at pH 4.4 and room temperature. In fact, under these experimental conditions, the CD trace lacks the characteristic local minima at 208/222 nm (Figure 2, spectrum a). Similar results were also observed at low temperatures (4 °C) and at different pH values (4–10) (results not shown). The spectrum was independent of protein concentration over a 10-fold concentration range (results not shown). By analogy to previous proposals [2], the N domain of TTF-1 may not be experiencing an aqueous environment in vivo because, as a transcriptional activator [12], it should be heavily involved in protein–protein interactions. Thus spectra were also collected in the presence of TFE to simulate a more hydrophobic milieu. The addition of TFE significantly alters the structure of the N domain. Increasing concentration of TFE leads to decrease in [θ] values at 208 and 222 nm, which is indicative of an increase in α-helical content (Figure 2). The amount of induced secondary structure reaches a plateau between 40 and 50 % TFE. The percentage of α-helix content induced by TFE, as well as the percentage of expected α-helix content (Figure 1, left-hand panel), are reported in Table 1. The results with the N domain are similar to those obtained with VP16, for which a gain of α-helix in 50 % TFE has been reported [6]. However, these results differ from those obtained on the acidic activation domains of the GCN4 and GAL4 proteins [8]. CD experiments in 50 % TFE at pH 5.5 showed that these domains fold in fact mostly as β-sheets.

Proteolysis of TTF-1 N domain

Limited proteolysis is a useful procedure to probe structural and dynamic aspects of globular proteins ([35] and references cited
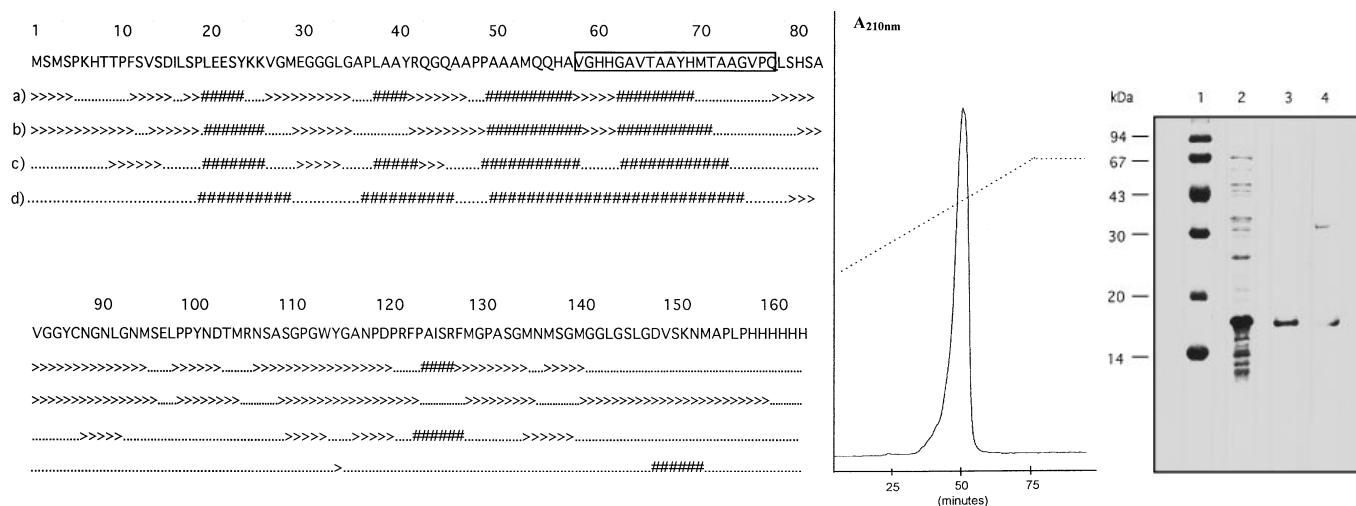
Figure 1   Purification of TTF-1 N domain

Left-hand panel: predicted primary structure of TTF-1 N domain. The numbering above the sequence refers to the residue positions. The cDNA coding for the amino acids 1–156 of rat TTF-1 was cloned into the vector pQ12 (Qiagen) in-frame with the coding region of six histidine residues. Thus the expressed protein (N domain–His$_6$) contains an extra hexahistidine sequence at the C-terminus, allowing for purification of the protein by Ni-NTA affinity chromatography. The N domain–His$_6$ fusion protein, consisting of 162 residues plus an initial methionine derived from the starting codon, was expressed in M15 *E. coli* cells. The symbols below the sequence correspond to secondary-structure predictions by methods of Nagano (a), Garnier–Osguthorpe–Robson (b), Chou–Fasman (c) and Rost–Sander (d). #, α-helix; >, β-turn. The boxed sequence corresponds to the fragment obtained after proteolysis with thermolysin in 10% TFE (see the Results and Discussion section). Middle panel: reverse-phase HPLC chromatography of the TTF-1 N domain on a C$_{18}$ column was performed after purification by Ni-NTA affinity chromatography to test for the achieved purity. A Pep RPC HR 5/5 (Pharmacia) column was employed and eluted at a flow rate of 0.8 ml/min with a linear gradient (————) of acetonitrile (50%, v/v, final concn.) in aq. 0.05% trifluoroacetic acid. Right-hand panel: SDS/12%-polyacrylamide gel, stained with Coomassie Blue, of TTF-1 N domain expression and purification by Ni-NTA affinity chromatography. Lane 1, molecular-mass markers (phosphorylase *b*, 94 kDa; albumin, 67 kDa; ovalbumin, 43 kDa; carbonic anhydrase, 30 kDa; trypsin inhibitor, 20.1 kDa; α-lactalbumin, 14.4 kDa); lane 2, *E. coli* crude lysate after induction with IPTG; lane 3, TTF-1 N domain after purification under reducing conditions [1 mM 2-mercaptoethanol (10 μg of protein)]; lane 4, TTF-1 N domain under non-reducing conditions (10 μg of protein).

Table 1   Comparison between α-helix content of TTF-1 N domain estimated from secondary predictions and CD spectra

Method	α-Helix (%)
Chou–Fasman	23[a]
Garnier–Osguthorpe	22
Nagano	22
Rost–Sander	30
Menendéz-Arias CD fitting	
50%TFE[b]	37.3 ± 4
0%TFE[c]	7.2 ± 3

[a] Percentages are expressed as summation of each structure versus contribution divided by 163 residues.
[b] Data for Menendéz-Arias calculation provided by Figure 2, spectrum f, from 250 to 197nm.
[c] Data for Menendéz-Arias calculation provided by Figure 2, spectrum a, from 250 to 197 nm.

therein). Globular proteins are cleaved at exposed and flexible loops only and never at chain segments of regular secondary structure [35]. In particular it has been reported that, when thermolysin, a TFE-resistant proteolytic enzyme, is used, the peptide-bond cleavage would occur at sites dictated by stereo-chemistry and flexibility of the polypeptide substrate and not by the specificity of the protease [35].

In order to further demonstrate the gain of secondary structure of the TTF-1 N domain in TFE solution, we performed a digestion with thermolysin in Tris buffer, pH 7.8. The SDS/
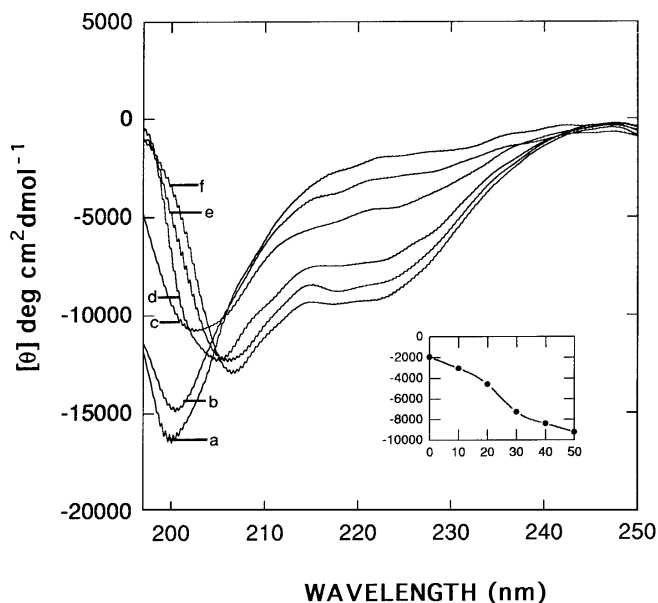
Figure 2   Far-UV CD spectra of the TTF-1 N domain

Spectra were acquired in 10 mM sodium acetate buffer, pH 4.4, at 25 °C. The protein concentration was 16 μM and the pathlength was 0.2 cm. Spectra (a–f) were obtained in the presence of (a) 0, (b) 10, (c) 20, (d) 30, (e) 40, (f) 50% (v/v) TFE respectively. The inset shows a titration curve reporting $[\theta]_{222}$ (in degrees·cm^2·dmol^{-1}) versus percentage (v/v) TFE. $[\theta]_{222}$ is indicative of α-helical structure.
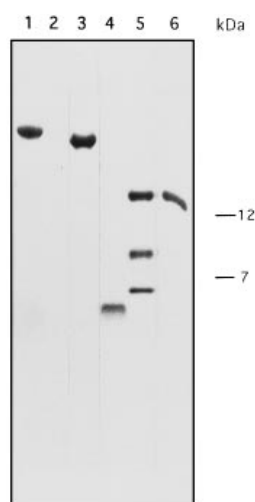
Figure 3   SDS/PAGE analysis of the proteolysis of TTF-1 N domain with thermolysin in aqueous TFE

The protein was digested at a protease/substrate ratio of 1:50 (w/w) in the presence of 0, 10 and 50 % (v/v) TFE. After 6 h reaction at 25 °C, an aliquot of the proteolysis mixture was analysed by SDS/PAGE. As a control, horse heart cytochrome c digestion was performed as described by Fontana et al. [35]. The Figure shows undigested TTF-1 N domain (lane 1), digested TTF-1 N domain in 0% TFE (lane 2), in 50% TFE (lane 3) and in 10% TFE (lane 4), undigested cytochrome c (lane 6) and digested cytochrome c in 50% TFE (lane 5).

PAGE analysis of the proteolysis mixture after 6 h reaction at 25 °C (Figure 3) reveals that, in the absence of TFE, the TTF-1 N domain is completely digested by thermolysin into small peptides which are not stained by Coomassie Blue (lane 2). In the presence of 50 % TFE, TTF-1 N domain is almost undigested (lane 3), confirming the evidence of gain of secondary structure (α-helix) obtained from CD studies. As a control of the functionality of thermolysin, limited proteolysis of cytochrome c was performed under the same conditions. In 50 % TFE, two fragments were observed (Figure 3, lane 5) according to the results previously reported by Fontana et al. [35].

As demonstrated by CD data, the TTF-1 N domain proved very sensitive, in gaining α-helical conformation, to TFE addition. The limited-proteolysis approach provides us with the opportunity to test whether the gain of α-helical structure observed at 10 % TFE is due to a structural transition of a discrete domain. Therefore the limited-proteolysis experiment was performed at 10 % TFE. As shown in Figure 3 (lane 4), at 10 % TFE, thermolysin is unable to digest completely the TTF-1 N domain, but it is possible to isolate a fragment protected from digestion. The fragment of approx. 2 kDa was eluted from gel and subjected to N-terminal protein sequencing. The sequence identifies a fragment of 20 residues between residues 58 and 78 (Figure 1, left-hand panel, boxed). Interestingly, this sequence maps inside the minimal activating region (residues 50–102) of TTF-1 N domain [12]. Therefore our results establish a link between sensitivity to α-helical-structure gain upon TFE addition and functional activity, suggesting the presence of a 'core domain' having a high propensity to experience unstructured $\rightleftharpoons$ structured transition upon contacts with other proteins of the transcription machinery. These data are in complete agreement with the 'induced fit' model proposed by Tjian and Maniatis [36] and confirmed by Schmitz et al. [7] and Dalhman-Wright and McEwan [37] for acidic domains.
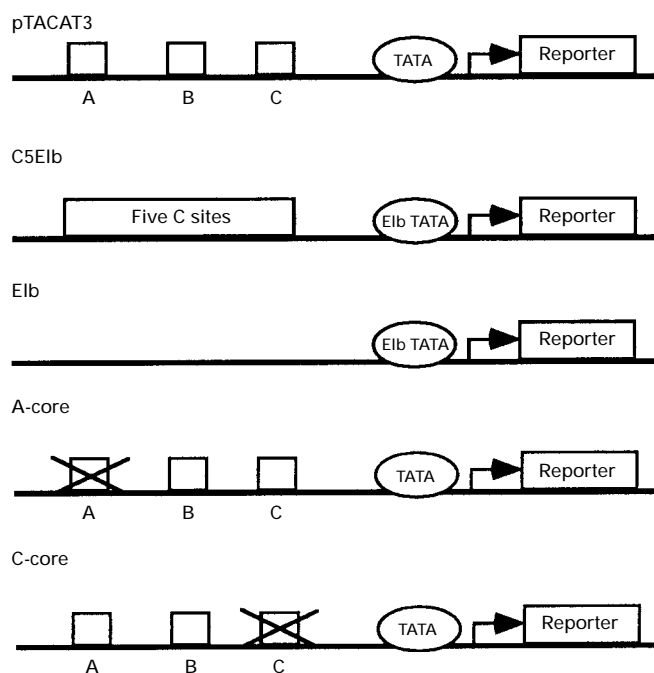
Figure 4   Structure of promoters used in HeLa transfection assays

pTACAT3 contains the native Tg promoter linked to the CAT reporter gene. The A, B and C boxes indicate TTF-1 binding sites [23]. C5E1b is an artificial promoter in which five C sites of Tg promoter have been inserted upstream of the E1b TATA box. E1b contains only the E1b TATA box in front of the CAT reporter gene. The A-core and C-core represent mutants of the Tg promoter. The deletion crosses indicate the presence of mutations that suppress TTF-1 binding to the respective site.

Functional studies in cell cultures

Previously obtained data demonstrate that a gain of secondary structure in hydrophobic conditions occurs with acidic activation domains [6]. Our data, based on CD and proteolytic studies, indicate that TFE-induced secondary structure also occurs in a transcriptional-activation domain that is not rich in acidic residues. Therefore the propensity to gain secondary structure under hydrophobic conditions is a phenomenon shared by different classes of transcriptional-activation domains.

The structural similarities among TTF-1 N domain and acidic activation domains call for further functional testing to complement and definitely establish an analogy. In particular, two functional expectations should be explicitly addressed: (i) acidic activation domains should be able to activate transcription in the same promoters where the N domain of TTF-1 is active; (ii) the TTF-1 N domain should compete with acidic activation domains for proteins of the basal transcriptional machinery (squelching effect [12]).

In order to test these expectations, a chimaeric molecule, containing the DNA-binding domain of TTF-1 linked to the acidic activation domain of VP16, was constructed (HDVP16). The transcriptional activation induced by this molecule was tested in contexts in which previous investigations have demonstrated that TTF-1 N domain is able to activate transcription [12]. Two different promoters were employed, namely the native Tg promoter (pTACAT3) and an artificial promoter (C5E1b) in which five C sites of the Tg promoter (recognized with high affinity by TTF-1) had been inserted upstream of the E1b TATA box [12] (Figure 4). Table 2 shows that either promoter is

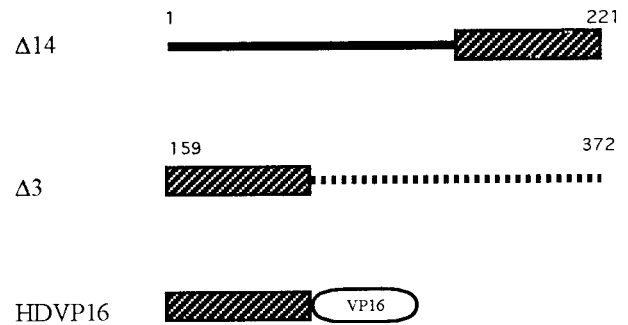| Activator	Promoter activity (CAT/luciferase activity ratio)				
	pTACAT-3	C5E1b	E1b	A-core	C-core
–	0.04 ± 0.01	0.05 ± 0.01	0.04 ± 0.02	0.09 ± 0.04	0.07 ± 0.03
HDVP16	3.93 ± 1.32	5.21 ± 1.65	0.05 ± 0.02	0.71 ± 0.39	0.40 ± 0.38

Table 3   The acidic domain of VP16 interferes with the function of TTF-1 N domain

In these experiments the reporter plasmid is always G5E1b which contains five GAL4-binding sites in front of the E1b TATA box and the activator used is always ΔG21, which is a chimaeric protein containing the TTF-1 N domain linked to the GAL4 DNA-binding domain. The structures of the competitors are shown schematically in Figure 5. Each value represents the mean ± S.D. for at least three independent experiments.

Competitor	Transcriptional activity (arbitrary units)
–	100
Δ3	82 ± 3
Δ14	0.4 ± 0.1
HDVP16	13.9 ± 1

Figure 5   Schematic structure of the competitors used in the experiment of Table 3

Thick black line, TTF-1 N domain; broken black line, TTF-1 C domain; hatched box, TTF-1 HD; empty box, VP16 trans-activation domain. Generation of the constructs is described in the Materials and methods section and in [12].

Table 4   Squelching effect by the TTF-1 N domain on the acidic activation domain of p65

In these experiments the reporter plasmid is always G5E1b, which contains five GAL4-binding sites in front of the E1b TATA box. G5E1b was introduced into HeLa cells together with the expression vectors and the CMV-Luc plasmid, expressing the luciferase gene (*luc*), and used as internal reference. HeLa cells were transfected with 5 μg of G5E1b, 0.5 μg of the activator GAL4-p65 and 5 μg of the competitor plasmid-expressing protein Δ14. Results are expressed as the CAT/luciferase activity ratio. Each value represents the mean ± S.D. for at least three independent experiments.

Activator	CAT/luciferase activity ratio
–	0.01 ± 0.01
Δ14	0.01 ± 0.01
GAL4-p65	0.23 ± 0.06
GAL4p65 + Δ14	0.06 ± 0.004

activated by the presence of HDVP16. The effect of HDVP16 is not observed in the basal promoter E1b (lacking the five C sites), indicating that the chimaeric molecule requires the binding to the DNA to exert its effect. These results indicate that the acidic activation domain of VP16 is able to activate transcription in the same contexts where the N domain of TTF-1 operates and provides us with a system to test whether restrictions affecting the activity of TTF-1 might also act on the chimaeric protein HDVP16. To this aim the transactivating function of HDVP16 on the A core and C core mutants of Tg promoter was tested. In these mutants the TTF-1 binding to either site A (A core) or C (C core) has been abolished (Figure 4). Previous studies had revealed that the TTF-1 activity with both mutant promoters was greatly reduced in either differentiated thyroid cells (FRTL-5) or HeLa cells expressing TTF-1 [12] compared with the wild-type Tg promoter. Table 2 shows that, with either mutant promoters, HDVP16 induces a transcriptional activity which is much lower than that observed with the wild-type Tg promoter. Thus, according to these data, on activating the Tg promoter, the acidic domain of VP16 is affected by the same regulatory constraints that operate with the wild-type TTF-1.

The expected competition between TTF-1 N domain and VP16 for common target(s) was confirmed by interference experiments (Table 3). The protein used as activator was ΔG21, which contains the N domain of TTF-1 linked to the DBD of GAL4 and is able to activate a promoter containing five GAL4 binding sites in front of the TATA box of E1b [12]. Proteins like HDVP16 and Δ14 (containing the TTF-1 N domain linked to the

HD; Figure 5), which possess the native TTF-1 DBD, but not that of GAL4, cannot recognize the GAL4-binding sites and could therefore interfere with the transcriptional activity of ΔG21 if they were capable of competing with the latter for components of the transcriptional machinery. Indeed, both Δ14, which contains the same transcriptional activation domain as ΔG21, and HDVP16 strongly inhibited the transcriptional activity of ΔG21. These results indicate that the activation domain of VP16 interacts with the same target(s) as the N domain of TTF-1. By contrast, when the competition in HeLa cells was established between ΔG21 and the protein Δ3, where the C domain of TTF-1 is appended to TTF-1 HD (Figure 5), no inhibition was observed, which suggests that the C domain of TTF-1 contacts target(s) different from those contacted by either VP16 or the N domain. The lack of interference by Δ3 is not due to low protein levels, since gel-retardation assays revealed that the proteins Δ14, HDVP16 and Δ3 were expressed at the same levels (results not shown).

This squelching effect was also demonstrated for another member of the acidic activation domains, namely the p65 subunit of the NF-κB transcription factor. This domain has been demonstrated to adopt α-helical conformation in TFE solution [7], suggesting that this phenomenon is not restricted to the only VP16 activation domain. As shown in Table 4, Δ14 is able to greatly reduce the transactivation activity induced by p65,
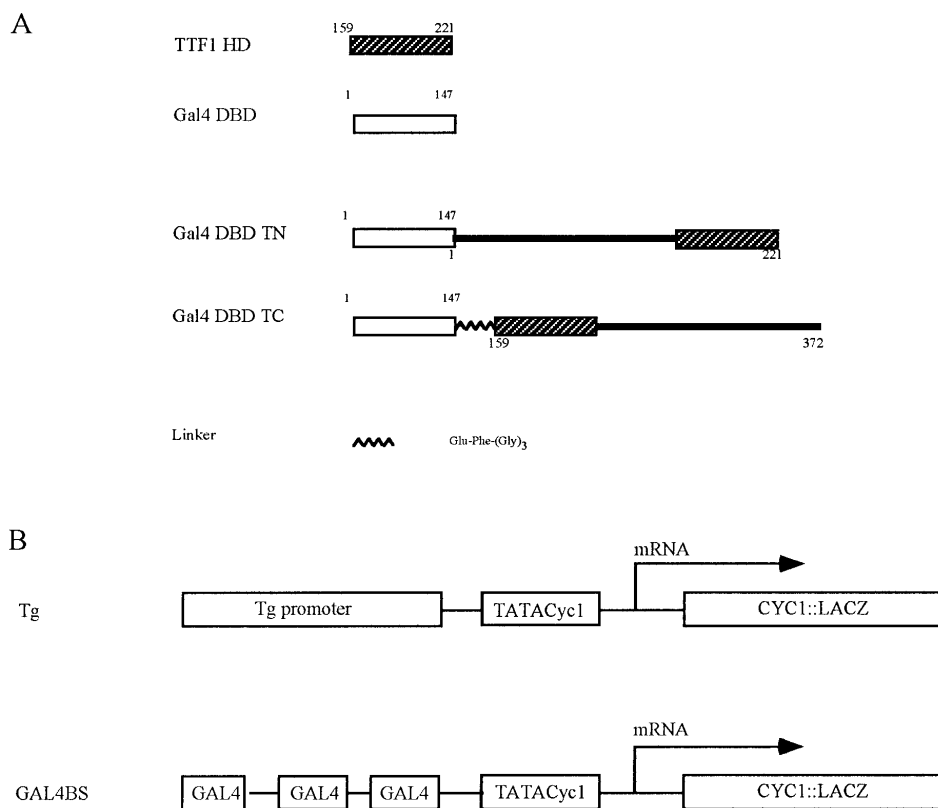
Figure 6   Schematic structure of chimaeric proteins and reporter genes used in *S. cerevisiae*

(**A**) Schematic view of the chimaeric proteins expressed in *S. cerevisiae*. In GAL4DBDTC the linker Glu-Phe-(Gly)$_3$ (shown as a thick wavy black line) was inserted as spacer between the GAL4DBD and the HD to prevent steric interferences between the two DNA-binding domains. (**B**) Schematic structure of the reporter genes.

suggesting that the N domain interacts with targets shared by the whole class of acidic activation domains.

## Functional studies in yeast *Saccharomyces cerevisiae*

The precise mechanisms by which the various classes of activation domains stimulate transcription are not known but, in the case of acidic activation domains, they appear to be conserved between yeast and higher eukaryotic cells. Only acidic activation domains are functionally active in different species [38]. For example, the mammalian activators VP16 and Jun oncoprotein efficiently activate transcription in *S. cerevisiae* [39,40]. Conversely the yeast activator GAL4 activates transcription also in *Drosophila* [41], tobacco (*Nicotiana tabacum*) [42] and human [43] cells. This species interchangeability is not present in the case of other classes of activation domains such as the glutamine-rich. In fact the activation domain of Sp1 fails to activate transcription in *S. cerevisiae* [44]. In proline-rich domains (e.g. AP-2, CTF/NF1) this allospecific activity depends upon the position of the promoter relative to the TATA box [45]. Taken together, these observations suggest that different classes of activators do interact with different parts of the basal eukaryotic transcriptional machinery and only mechanisms adopted by acidic domains are conserved from yeast to man. Therefore, if the TTF-1 N domain uses mechanisms similar to those used by acidic domains, it should be able to activate transcription in yeast. Two yeast expression vectors, GAL4DBDTN and GAL4DBDTC, were constructed that contained the N and C activation domains of TTF-1 respectively. The detailed structures of the expressed

Table 5   Transcriptional activity of chimaeric proteins on Tg and GAL4BS elements in *S. cerevisiae*

The values of transcriptional activity are obtained by means of a β-galactosidase assay of yeast liquid cultures [27]. Each value represents the mean $\pm$ S.D. for at least three independent experiments. TC, TTF-1 C domain; TN, TFF-1 N domain.

Activator	Transcriptional activity (arbitrary units)	
	Tg	GAL4BS
–	0.36 $\pm$ 0.1	0.15 $\pm$ 0.1
GAL4DBD	0.36 $\pm$ 0.1	0.25 $\pm$ 0.1
GAL4DBDTC	2.5 $\pm$ 0.5	1.5 $\pm$ 0.3
GAL4DBDTN	41.12 $\pm$ 2	20.15 $\pm$ 2.1

proteins are shown in schematic form in Figure 6(A). The transcriptional activity of these chimaeric proteins was tested with two different reporters: one containing the TTF-1 binding region of Tg promoter fused 5′ to the CYC1 TATA box (Tg in Figure 6B) and the other containing three GAL4 binding sites fused 5′ to the CYC1 TATA box (GAL4BS in Figure 6B). As shown in Table 5, GAL4DBDTN, which contains the TTF-1 N domain, is able to efficiently activate transcription in both the Tg and GAL4BS contexts. In contrast, the GAL4DBDTC, which contains the TTF-1 C domain, activates transcription very weakly, only two- to three-fold with respect to the basal values.
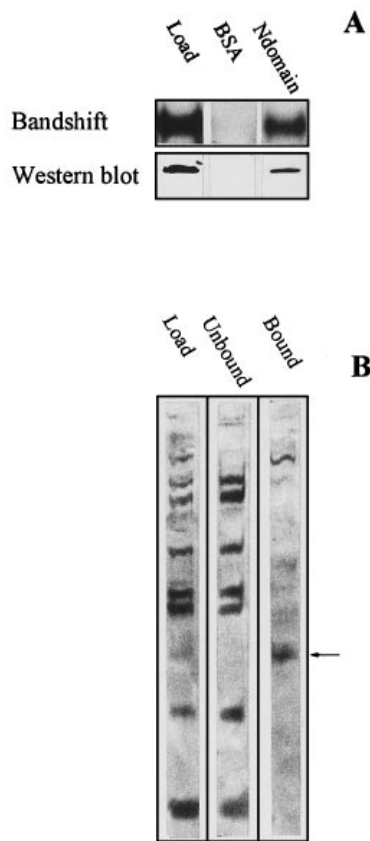
Figure 7   Recombinant yeast TBP binds directly to TTF-1 N domain

(**A**) Eluted fractions from precipitation with BSA (central lane) or TTF-1 N domain (right lane) linked to CNBr-activated Sepharose 4B resin assayed either in bandshift assay (top) or in Western blot (bottom). See the Materials and methods section for details. (**B**) SDS/12%-PAGE gel of partially purified recombinant TBP input fraction (left lane), unbound fraction (central lane) and eluted fraction (right lane) from precipitation with TTF-1 N domain linked to CNBr-activated Sepharose 4B resin. The arrow indicates the recombinant yeast TBP polypeptide. The gel was stained with silver.

The transcriptional activity obtained by the TTF-1 N domain is only twofold lower than that achieved by the wild-type GAL4 protein (results not shown). Therefore these data indicate that the TTF-1 N domain is able to efficiently activate transcription in yeast, a property typical of acidic activation domains.

Interaction with yeast TBP

A typical event in transcriptional stimulation by acidic activators is the binding to TBP. To assess whether the N domain uses mechanisms similar to those used by acidic activators, we assayed the ability of the N domain to interact *in vitro* with the pivotal element TBP. We expressed yeast TBP in *E. coli* and partially purified on a Mono-S column reaching a degree of purification of about 10%. This relatively low degree of purification was useful to assay the specificity of interaction. The *in vitro* interaction was tested by precipitation experiments using the N domain coupled to CNBr-activated Sepharose 4B beads. As a control we used BSA coupled to the same resin. After incubating 250 μl of partially purified TBP, the beads were extensively washed with binding buffer (see the Materials and methods section) and eluted with 1 M KCl. The eluted fractions were assayed for the presence of TBP either in bandshift assays or in

Western blot. As shown in Figure 7(A), only the N domain was able to retain TBP. This interaction is specific, as confirmed by the SDS/PAGE shown in Figure 7(B), which demonstrates that only the TBP band (arrow) is enriched after elution. These data demonstrate that, similarly to the acidic domains, TTF-1 N domain directly interacts with the TATA-binding protein, strongly suggesting that both domains, though having no sequence similarity at all, use similar pathways to activate transcriptional machinery.

A central issue for HD-containing proteins and, more broadly, for promoter-specific transcription factors, is the understanding of the molecular mechanisms by which these proteins achieve distinct functional specifications. Although several experiments with HD-containing proteins from *Drosophila* have shown that most of biological specificity is due to the HD [46,47], results that challenge this view have also been reported [48]. In our functional experiments, the only domain shared by HDVP16 and the wild-type TTF-1 is the DBD (HD). Moreover, the TTF-1 N domain possesses most of the biochemical properties of acidic activation domains. Therefore our data would emphasize the critical role that the DBDs of transcription factors play in the specification of the activity of these proteins. Accordingly, Wang et al. [49] demonstrated, in a knock-in experiment, that the two related transcription factors, Myf5 and myogenin, belonging to the bHLH family, though having different transcriptional activation domains, are functionally interchangeable and redundant in rib formation. Recent data on Pax-6 transcription factors [50] reinforce this view. In fact, targeted expression in *Drosophila* imaginal discs of either mouse, squid or *Drosophila* protein results in supernumerary eyes. Pax-6 proteins are very much conserved in the DBDs (both HD and paired domains), but are quite divergent in the activation domain (PST domain).

Our data would predict that an acidic activation domain should be able to substitute the N domain in the context of TTF-1. A similar prediction has been successfully verified in *Drosophila* in the context of the bicoid protein [47]. Knock-out experiments revealed that the deletion of the TTF-1 gene induces dramatic phenotypic alterations in different tissues (brain, lung and thyroid) [51]. Knock-in experiments with chimaeric genes should provide an excellent tool to address the issue for mammalian proteins.

Activation domains are grouped on the basis of the chemical properties of their residues, particularly the richness in particular amino acids. This kind of classification is supported by findings indicating that, in some situations, the specificity of an activation domain is affected by an uneven composition of its sequence. For instance, glutamine-rich domains are unable to function in yeast because of the lack, in this organism, of the specific co-activators [38]. However, several groups have recently demonstrated that specific bulky hydrophobic amino acids, rather than acidic residues, are important for the function of several activation domains ([5,52] and references cited therein). Our results indicate that a non-acidic domain may activate transcription by using molecular mechanisms similar to those used by acidic domains. Therefore we propose that grouping the activation domains on the basis of the chemical properties of their component residues could have only a limited functional relevance. By contrast, we identify a class of transcriptional activators which include the TTF-1 N domain and acidic domains. Members of this class possess similar structural features and activate transcription through mechanisms which are highly conserved with the evolution of eukaryotic cells.

REFERENCES

1 Darnell, J. E. (1982) Nature (London) **297**, 365–371

2 Mitchell, P. J. and Tjian, R. (1989) Science **245**, 371–378

3 Pabo, C. O. and Sauer, R. T. (1992) Annu. Rev. Biochem. **61**, 1053–1095

4 Triezenberg, S. J. (1995) Curr. Opin. Genet. Dev. **5**, 190–196

5 Regier, J.L, Shen, F. and Triezenberg, S. J. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 883–887

6 Donaldson, L. and Capone, J. P. (1992) J. Biol. Chem. **267**, 1411–1414

7 Schmitz, M. L., dos Santos Silva, M. A., Altmann, H., Czisch, M., Holak, T. A. and Baeuerle, P. A. (1994) J. Biol. Chem. **269**, 25613–25620

8 Van Hoy, M., Leuther, K. K., Kodadek, T. and Johnston, S. A. (1993) Cell **72**, 587–594

9 Damante, G. and Di Lauro, R. (1994) Biochim. Biophys. Acta **1218**, 255–266

10 Musti, A. M., Ursini, V. M., Avvedimento, V. E., Zimarino, V. and Di Lauro, R. (1987) Nucleic Acids Res. **15**, 8149–8166

11 Bruno, M. D., Bohinski, R. J., Huelsman, K. M., Whitsett, J. A. and Korfhagen, T. R. (1995) J. Biol. Chem. **270**, 6531–6536

12 De Felice, M., Damante, G., Zannini, M., Francis-Lang, H. and Di Lauro, R. (1995) J. Biol. Chem. **270**, 26649–26656

13 Wetlaufer, D. B. (1962) Adv. Protein Chem. **17**, 303–390

14 Lowry, O. H., Rosebrough, N. J., Farr, A. L. and Randall, R. J. (1951) J. Biol. Chem. **193**, 265–275

15 Bradford, M. (1976) Anal. Biochem. **72**, 248–254

16 Petri, V., Hsieh, M. and Brenowitz, M. (1995) Biochemistry **34**, 9977–9984

17 Menendéz-Arias, L., Gomez-Gutiérrez, J., Garcia-Fernandez, M., Garcia-Tejedor, A. and Moran, F. (1989) Comput. Appl. Biosci. **4**, 479–482

18 Shagger, H. and von Jagow, G. (1987) Anal. Biochem. **166**, 368–379

19 Grahm, F. L. and van der Eb, A. J. (1973) Virology **52**, 456–467

20 Gorman, C. M., Moffat, L. M. and Howard, B. H. (1982) Mol. Cell. Biol. **2**, 1044–1051

21 de Wet, J. R., Wood, K. V., Deluca, M., Helinski, D. R. and Subramani, S. (1987) Mol. Cell. Biol. **7**, 725–737

22 Sinclair, A. J., Lonigro, R., Civitareale, D., Ghibelli, L. and Di Lauro, R. (1990) Eur. J. Biochem. **193**, 311–318

23 Civitareale, D., Lonigro, R., Sinclair, A. J. and Di Lauro, R. (1989) EMBO J. **8**, 2537–2542

24 Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. and Pease, L. R. (1989) Gene **77**, 51–60

25 Sherman, F., Fink, G. R. and Hick, J. (1986) Methods in Yeast Genetics, Cold Spring Harbor Laboratory Press, Plainview, NY

26 Schiestl, R. H. and Gietz, R. D. (1989) Curr. Genet. **16**, 339–346

27 Guarente, L. (1983) Methods Enzymol. **101**, 181–191

28 Kim, Y., Geiger, J. H., Hahn, S. and Sigler, P. B. (1993) Nature (London) **365**, 512–520

29 Hamodrakas, S. J., Jones, C. W. and Kafatos, F. C. (1982) Biochim. Biophys. Acta **700**, 42–51

30 Chou, P. Y. and Fasman, U. D. (1974) Biochemistry **13**, 211–215

31 Garnier, L., Osguthorpe, D. J. and Robson, B. (1978) J. Mol. Biol. **120**, 97–120

32 Nagano, K. and Hashegawa, K. (1975) J. Mol. Biol. **75**, 401–420

33 Rost, B. and Sander, C. (1993) J. Mol. Biol. **232**, 584–599

34 Arnone, M. I., Zannini, M. and Di Lauro, R. (1995) J. Biol. Chem. **270**, 12048–12055

35 Fontana, A., Zambonin, M., De Filippis, V., Bosco, M. and Polverino de Laureto, P. (1995) FEBS Lett. **362**, 266–270

36 Tjian, R. and Maniatis, T. (1994) Cell **77**, 5–8

37 Dahlman-Wright, K. and McEwan, I. J. (1995) Biochemistry **35**, 1323–1327

38 Sauer, F. and Tjian, R. (1997) Curr. Opin. Genet. Dev. **7**, 176–181

39 Chasman, D. I., Leatherwood, J., Carey, M., Ptashne, M. and Kornberg, R. D. (1989) Mol. Cell. Biol. **9**, 4746–4749

40 Struhl, K. (1988) Nature (London) **332**, 649–650

41 Fischer, J. A., Giniger, E., Maniatis, T. and Ptashne, M. (1988) Nature (London) **332**, 853–856

42 Ma, J., Przibilla, E., Hu, J., Bogorad, L. and Ptashne, M. (1988) Nature (London) **334**, 631–633

43 Webster, N., Jin, J. R., Green, S., Hollis, M. and Chambon, M. (1988) Cell **52**, 169–178

44 Ponticelli, A. S., Pardee, T. S. and Struhl, K. (1995) Mol. Cell. Biol. **15**, 983–988

45 Kunzler, M., Braus, G. H., Georgiev, O., Seipel, K. and Schaffner, W. (1994) EMBO J. **13**, 641–645

46 Gibson, G., Schier, A., LeMotte, P. and Gehring, W. J. (1990) Cell **62**, 1087–1103

47 Driever, W., Ma, J., Nusslein-Volhard, C. and Ptashne, M. (1989) Nature (London) **342**, 149–154

48 Chan, S. K. and Mann, R. S. (1993) Genes Dev. **7**, 796–811

49 Wang, Y., Schnegelsberg, J. D. and Jaenisch, R. (1996) Nature (London) **379**, 823–825

50 Tomarev, S. I., Callaerts, P., Kos, L., Zinovieva, R., Halder, G., Gehring, W. and Piatigorsky, J. (1997) Proc. Natl. Acad. Sci. U.S.A. **94**, 2421–2426

51 Kimura, S., Hara, Y., Pineau, T., Fernandez-Salguero, P., Fox, C. H., Ward, J. M. and Gonzalez, F. J. (1996) Genes Dev. **10**, 60–69

52 Sainz, M. B., Goff, S. A. and Chandler, V. I. (1997) Mol. Cell. Biol. **17**, 115–122