Recruitment of the transcriptional machinery through GAL11P: structure and interactions of the GAL4 dimerization domain

Patricia Hidalgo,1,2,4 Aseem Z. Ansari,3,6 Peter Schmidt,1 Brian Hare,1 Natasha Simkovich,3 Susan Farrell,3,5 Eun Ji Shin,3 Mark Ptashne,3,6 and Gerhard Wagner1,7

1Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA; 2Bristol-Myers Squibb Pharmaceutical Research Institute fellow of the Life Sciences Research Foundation, Princeton, New Jersey 08544; 3Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts 02138, USA

The GAL4 dimerization domain (GAL4-dd) is a powerful transcriptional activator when tethered to DNA in a cell bearing a mutant of the GAL11 protein, named GAL11P. GAL11P (like GAL11) is a component of the RNA–polymerase II holoenzyme. Nuclear magnetic resonance (NMR) studies of GAL4-dd revealed an elongated dimer structure with C2 symmetry containing three helices that mediate dimerization via coiled-coil contacts. The two loops between the three coiled coils form mobile bulges causing a variation of twist angles between the helix pairs. Chemical shift perturbation analysis mapped the GAL11P-binding site to the C-terminal helix H9251 and the loop between H92511 and H92512. One GAL11P monomer binds to one GAL4-dd dimer rendering the dimer asymmetric and implying an extreme negative cooperativity mechanism. Alanine-scanning mutagenesis of GAL4-dd showed that the NMR-derived GAL11P-binding face is crucial for the novel transcriptional activating function of the GAL4-dd on GAL11P interaction. The binding of GAL4 to GAL11P, although an artificial interaction, represents a unique structural motif for an activating region capable of binding to a single target to effect gene expression.

[Key Words: GAL4; dimerization domain; GAL11P; NMR structure]

Received December 14, 2000; revised version accepted February 20, 2001.

The GAL4 transcriptional activator is required for the regulation of genes involved in galactose and melibiose metabolism in the yeast Saccharomyces cerevisiae (Johnston 1987). GAL4 has a modular structure in which distinct regions of the molecule mediate recognition and binding to DNA, activation of transcription, and dimerization (Fig. 1A; Laughon 1984; Laughon et al. 1984; Keegan et al. 1986; Ma and Ptashne 1987; Lin et al. 1988; Carey et al. 1990). Only the structure of the DNA-binding domain (residues 1–66) has been solved alone in solution [Baleja et al. 1992, Kraulis et al. 1992] and in complex with DNA [Marmorstein et al. 1992]. It contains an N-terminal DNA-recognition domain (residues 7–40), in which six cysteines bind two Zn+2 ions. In the free protein in solution, the region from residue 41–66 is unstructured. In the crystal structure of the complex with DNA, residues 50–63 form a coiled-coil dimerization element that extends perpendicularly away from the DNA helix [Marmorstein et al. 1992]. Nevertheless, GAL4 1–65 is a monomer in the absence of DNA. Functional characterization of the 881 residue GAL4 protein has shown that a DNA-independent dimerization domain extending beyond residue 63 comprising the sequence from 65 to 94 [Carey et al. 1989]; however, no structure of a construct containing the complete dimerization domain has been reported yet. Transcriptional activation function in yeast has been assigned to two regions: the primary activating region comprising residues 768–881 and a weaker one that lies between residues 148–196 [Keegan et al. 1986; Ma and Ptashne 1987; see Fig. 1]. GAL450–97ordinarily mediates dimerization (Carey et al. 1989). However, the same fragment manifests a novel transcriptional activating capability in yeast cells carrying a single point mutation in the GAL11 protein. This mutant GAL11P (P standing for transcriptional potentiator) and wild-type GAL11 are located in an ∼2-MD complex with RNA polymerase and 50 or more polypeptides in what has been called the RNA–polymerase II holoen-
zyme (Kim et al. 1994; Barberis et al. 1995; Hengartner et al. 1995).

The P mutation replaces the N342 residue with one of any number of hydrophobic amino acids and confers to GAL11 the ability to bind to the GAL4 dimerization domain (GAL4-dd). On the contrary, GAL4-dd does not interact in vitro with wild-type GAL11 as detected by several methods (see below). Furthermore, the strength of the interaction between the GAL4-dd and each GAL11P allele as measured in vitro correlates directly with the level of gene activation observed in vivo. The interacting regions of GAL4 and GAL11P can be exchanged such that the P region of GAL11P is tethered to DNA, and the GAL4-dd is attached to the holoenzyme without loss of transcriptional activation. Taken together, these results strongly suggest that recruitment of the holoenzyme to the promoter, whether by the fortuitous GAL4–GAL11P interaction or by a classic activator, results in transcriptional activation (Barberis et al. 1995; Simkovich 1996). The C-terminal region of GAL4 (851–881) recognizes the GAL80 inhibitory protein (not shown). [B] Partial amino acid sequences of the GAL4-dd and LAC9. Conserved residues are shaded in yellow. The double mutation on GAL4-dd (87 and 90) are colored red (bold). The secondary structure elements, a1, a2, and a3 are indicated by blue cylinders. [C] Temperature dependence of the unfolding of GAL4-dd. The denaturation and renaturation curves are shown in pink and blue, respectively. The CD spectra were taken in a sample 0.5 mg/mL in 50 mM phosphate buffer at pH 7.2 in an AVIV instrument, model 62DS (AVIV Associates) at 222 nm.

Results

GAL4-dd adopts a dynamic state lacking a sharp thermal unfolding transition

To find the best conditions for a structure determination, we recorded 1H–15N heteronuclear single quantum coherence (HSQC) spectra of GAL4-dd over a range of pH
and temperature. The spectra were best in a narrow range of pH near 7.5 and at 35°C. Variation of the temperature between 10°C and 50°C results in large spectral changes that are different from the kind usually observed for cooperative thermal unfolding. Although several sharp signals that were later assigned to the C-terminal tail stayed almost invariant, most other signals changed line shapes with temperature, disappeared, or newly appeared in an uncorrelated manner. This indicates temperature-dependent noncooperative conformational changes. Consistently, the temperature dependence of the circular dichroism (CD) spectrum shown in Figure 1C lacks a sharp cooperative unfolding transition and resembles the flat melting curves of helix-coil transitions in short helical polypeptides [Zimm and Bragg 1959]. Aware of the dynamic state of the protein, we pursued a characterization of GAL4-dd at 35°C and pH 7.4.

GAL4-dd contains three helices that dimerize via coiled-coil interactions

Figure 1B shows the amino acid sequence of the GAL4-dd [50–106] aligned with the primary sequence of Lac9, the transcriptional activator that regulates the expression of galactose and lactose genes in Kluyveromyces lactis [Das and Hollenberg 1982; Salmeron and Johnston 1986; Wray et al. 1987]. Resonance assignments of GAL4-dd were pursued with NMR triple-resonance and nuclear Overhauser enhancement (NOE) methods. Of the 57 residues, 43 were assigned unambiguously. Five residues at the N terminus [50–54] and four residues at the C terminus [101 and 103–105] could not be assigned. Flexible tails often remain unassigned in NMR analyses. Furthermore, nine residues from interior sites, subsequently showed to be located in connecting loop regions [67, 69, 73–77, 84, and 86], could not be identified. Either they were absent or the few unassigned signals showed no spectral properties enabling assignment. The absence of peaks is consistent with the dynamic state of GAL4-dd mentioned above.

Observation of characteristic NOE patterns and analysis of chemical shifts revealed that there are three short helices at positions 55–66, 78–84, and 87–97 [Fig. 1B]. However, there are no long-range contacts between the three helices indicating that the structure is extended. The first helix α1 coincides with the terminal helix in the crystal structure of the complex of GAL41–65 with DNA [Marmorstein et al. 1992].

To characterize the mode of dimerization, we followed a previously described procedure [Walters et al. 1997a] in which a uniformly 15N- and 2H-labeled sample was mixed with unlabeled protein. In H2O solution, all amide groups are protonated, and all cross peaks between 15N-bound amide protons and aliphatic protons in a three-dimensional 15N-dispersed NOE spectroscopy [NOESY] experiment must be across the dimerization interface. In this way, 40 intermonomer NOE contacts were found (involving 16 residues of each monomer), 24 for α1, two for the loop between α1 and α2, six for α2, six for α3, and two for the region beyond α3.

Structure of the GAL4 dimerization domain

Structures of GAL4-dd were calculated with a simulated annealing protocol within the CNS software package [Brünger et al. 1998]. The constraints used for the structure calculations are listed in Table 1. They consist of short and medium-range intermonomer constraints that establish the secondary structure of the three helices and the 40 intermonomer constraints that determine the parallel dimer structure. Because only one resonance signal per residue was observed, the dimer must be symmetric, and C2 symmetry was enforced in the structure calculations. Figure 2A shows a ribbon diagram of a representative structure. The overall dimensions are −13 Å × 62 Å × 5 Å. Figure 2B shows the backbone atoms of the 17 lowest energy structures with the C, Cα, and N atoms of α1, α2, and α3 simultaneously superimposed. The average root mean-square deviation (rmsd) from the mean structure is 3.8 Å. However, the rmsd values decrease considerably when the backbone atoms of the helices are separately superimposed and become 0.45 Å, 1.7 Å, and 0.75 Å for α1 [residues 55–66], α2 [residues 78–84], and α3 [residues 87–97], respectively.

Table 1. Structural statistics for the NMR structure calculations

<table>
<thead>
<tr>
<th>Restraint type</th>
<th>Number of restraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOE distance restraints per monomer</td>
<td>436</td>
</tr>
<tr>
<td>(TOTAL)</td>
<td></td>
</tr>
<tr>
<td>Intra-residue</td>
<td>230</td>
</tr>
<tr>
<td>Inter-residue</td>
<td>206</td>
</tr>
<tr>
<td>Sequential</td>
<td>113</td>
</tr>
<tr>
<td>Medium</td>
<td>93</td>
</tr>
<tr>
<td>i,i + 2</td>
<td>44</td>
</tr>
<tr>
<td>i,i + 3</td>
<td>40</td>
</tr>
<tr>
<td>i,i + 4</td>
<td>9</td>
</tr>
<tr>
<td>Hydrogen Bonds</td>
<td>38</td>
</tr>
<tr>
<td>Inter-monomer</td>
<td>40</td>
</tr>
<tr>
<td>Ramachandran plot for all assigned residues</td>
<td></td>
</tr>
<tr>
<td>Most favorable region (%)</td>
<td>71.0</td>
</tr>
<tr>
<td>Additionally allowed region (%)</td>
<td>20.7</td>
</tr>
<tr>
<td>Generously allowed region (%)</td>
<td>5.1</td>
</tr>
<tr>
<td>Disallowed region (%)</td>
<td>3.2</td>
</tr>
<tr>
<td>Average RMSD from idealize covalent geometry</td>
<td></td>
</tr>
<tr>
<td>Bonds (Å)</td>
<td>0.002 ± 0.0001</td>
</tr>
<tr>
<td>Angles (°)</td>
<td>0.420 ± 0.0140</td>
</tr>
<tr>
<td>Improper s</td>
<td>0.292 ± 0.0234</td>
</tr>
<tr>
<td>Average RMSD (Å) of C, Cα, N backbone atoms to the average structure</td>
<td></td>
</tr>
<tr>
<td>α1, α2, α3</td>
<td>3.8</td>
</tr>
<tr>
<td>α1 [residues 55–66]</td>
<td>0.45</td>
</tr>
<tr>
<td>α2 [residues 78–84]</td>
<td>1.70</td>
</tr>
<tr>
<td>α3 [residues 87–97]</td>
<td>0.75</td>
</tr>
</tbody>
</table>

No violation greater than 0.3 Å was found. The 18 residues with unassigned amide group including proline 73 are: 50–54, 67, 69, 74–77, 84, 86, 101, and 103–105.

Average RMSD (Å) of C, Cα, N backbone atoms to the average structure

α1, α2, α3 3.8
α1 [residues 55–66] 0.45
α2 [residues 78–84] 1.70
α3 [residues 87–97] 0.75
Hidalgo et al.

α3 (residues 87–97), respectively (Fig. 2C). α1 and α3 helices show the classic coiled-coil conformation. The α3 helices are less super-coiled than the α1 pair, however, and their interactions seems to be weaker as fewer inter-helix contacts were observed. The short α2 helices pack onto each other with angles >90° in an almost antiparallel orientation (see Fig. 2A). A relatively long loop connects α1 and α2, whereas α2 and α3 are linked by

**Figure 2.** Three-dimensional structure of the GAL4-dd. (A) A stereoview of the ribbon diagram of a representative structure of the GAL4-dd showing residues H53 to D100. The N and C termini are labeled N and C, respectively. The two monomers are colored differently. Unassigned residues P73 to L77 within the loop connecting α1 and α2 are colored orange. This figure was generated with the program MOLMOL [Koradi et al. 1996]. (B) A stereoview of the backbone atoms from residues L54 to D100 of the 17 lowest energy structures with helices α1, α2, and α3 (blue) simultaneously superimposed. The average backbone RMSD of α1, α2, and α3 to the mean structure is 3.8 Å. (C) A stereoview of the 17 lowest energy structures whereby the backbone atoms from residues H53 to D100 of each of the three helices were separately superimposed. For clarity, the loops regions are displayed for only one structure. The backbone atoms of the unassigned P73 to L77 segment are colored in orange as in A. The average backbone RMSD of α1, α2, and α3 to the mean structure is 0.45, 1.7, and 0.75 Å, respectively. (D) Ensemble of the 17 structures aligned by superimposing the backbone heavy atoms of helix α1.
only a couple of residues. The longest loop bears a strictly hydrophobic segment from F68 to F72. In contrast, the subsequent segment from P73 to L77 is predominantly hydrophilic. Because there was no observable signal for any of the latter residues, we conclude that this segment might undergo slow conformational exchange.

The disorder between the helices, which is apparent in Figure 2B, is primarily a variation of the amount of twist around the C$_2$ symmetry axis between the a1–a1', o2–o2', and o3–o3' coiled coils. This is shown in Figure 2D in which the structures were aligned by superimposing only the backbone atoms of the a1 helices. The figure shows that there is variability of orientation around the C$_2$ symmetry axis, of the large loop, the o2 and o3 helices relative to the a1–a1'. The tilt angle between the a1–a1' and o3–o3' coiled coils varies rather little in the ensemble of calculated structures. This is reasonable because any permanent tilt would make the dimer asymmetric, which would contradict the symmetric nature of the observed NMR spectra. From these considerations, we conclude that the average solution structure of GAL4-dd is extended as shown in Figure 2A, and the disorder consists of the amount of twist and presumably a transient variation of the lengths of the helices.

Dimerization interface

The GAL4-dd dimerizes in solution because of multiple coiled-coil type interactions between symmetry related atoms situated along each monomer. The identification of the intermonomers contacts forming the dimer interface was essential for obtaining the structure of the GAL4 dimer. Intermonomer NOEs were obtained using a previously described asymmetric deuteration strategy [WALTERS et al. 1997a].

Figure 3A and B shows the superposition of the 17 lowest energy structures with the side chains of residues forming the dimer interface within a1 and o3 displayed. A projection of the helices onto a wheel diagram is shown in Figure 3C and D. Annotating the heptad repeat sequence characteristic of coiled coils [LUPAS 1996] by letters from a to g, a1 contains the sequence a$_1$b$_1$c$_1$d$_1$e$_1$f$_1$g$_1$a$_2$b$_2$c$_2$d$_2$ in which the hydrophobic residues (in boldface) at positions a$_1$, d$_1$, a$_2$, and d$_2$ corresponding to the residues L54, V57, L61, and L64, respectively, form the core of the dimer interface. The same arrangement was reported for a1 in the crystal structure of GAL4 complexed with DNA [MARMORSTEIN et al. 1992]. The disruption of the coiled-coil conformation occurs at the second half of the second heptad [c$_3$f$_3$g$_3$] where position g$_3$ is occupied by a hydrophobic residue [L67] instead of a hydrophilic one presumably causing the interruption of the coiled coil.

Among the residues located in the loop between a1 and o2, P71 shows intermonomer contacts. In the short helix o2, I80, L81, and K82 also show dimer contacts. Finally, in o3, we observe intermonomers contacts for I89, E90, and L93, but the packing of the o3 helices appears to be less tight than in a1. Considering that the number of observed intermonomer and intramonomer NOE distance restraints was higher for a1 coiled coil, the differences may reflect variations in the dynamic behavior of the molecule. Therefore, we assayed this possibility by studying the mobility of the GAL4-dd. The dynamic properties of the GAL4-dd may be relevant for the transcriptional activation function as will be discussed below.

Protein mobility

To understand whether the apparent disorder in some parts of the GAL4-dd structure is due to mobility rather than lack of constraints, and to provide a basis for a possible correlation between mobility and activation function, we studied the motional properties of the GAL4-dd by using $^{15}$N relaxation experiments. Three relaxation parameters were measured, the nitrogen longitudinal relaxation rate $R_{1N}(N_x)$, the nitrogen transverse relaxation rate $R_{2N}(N_{xy})$, and the heteronuclear NOE (XNOE) [Fig. 4; PENG and WAGNER 1992]. The $^{15}$N relaxation parameters are related to the rotational diffusion of the N–H bond due to overall tumbling and internal motion, and so they contain information on the molecular mobility of
Increased protein mobility is manifested in faster $R_N(N_z)$ rates, slower $R_N(N_{x,y})$ rates, and augmented XNOE magnitudes. The GAL4-dd (55–102) displays average values for these relaxation parameters comparable to those measured for related proteins of similar size, such as PUT3 (Walters et al. 1997b). The average values are 2.43/sec for $R_N(N_z)$ with a range of 1.66 to 2.88/sec, 9.48/sec for $R_N(N_{x,y})$ with a range of 4.48 to 12.9/sec, and 0.48 for the XNOE with a range of 0.22 to 1.94. It is readily apparent from Figure 4 that the relaxation behavior of the GAL4-dd is not uniform throughout its amino acid sequence. The N-terminal half bearing a displays a lower average mobility than the C-terminal half composed of $\alpha_2$ and $\alpha_3$ with an overall average $R_N(N_{x,y})$ of 10.6/sec and 9.3/sec and an overall XNOE average value of 0.32 and 0.42, respectively. Leu 61 and Leu 64, which form the core of the dimer interface in $\alpha_1$, show the lowest XNOE values. Of note, the mobility of $\alpha_3$ increases continuously toward the end and continues to grow toward the unstructured end (Fig. 4B,C). This is consistent with the picture of a fracting helix with a poorly defined helix boundary.

The different dynamics properties between the $\alpha_1$ and $\alpha_2$–$\alpha_3$ regions can explain, partly, the relatively reduced number of NOEs observed for the C-terminal region of the GAL4-dd and consequently, the better structural definition for $\alpha_1$ than $\alpha_2$–$\alpha_3$. The number of intermonomer contacts observed for each residue is indicated in Figure 4C.

**Alanine-scanning mutagenesis to probe the GAL4-dd–GAL11P261–352 interaction**

**Activation in vivo** Previous truncation analyses identified the GAL4 residues 58–97 as the minimal segment necessary for interaction in vivo and electrophoretic mobility shift assays identify GAL4 mutants that bind GAL11P263–352. A] In vivo activation assay. (Top) Design of the experiment: Two LexA-binding sites are located 50 bp upstream of the GAL1 promoter TATA box. The expression of the reporter gene (lacZ) is regulated by the ability of LexA derivatives to interact with GAL11P and recruit the transcriptional machinery to this promoter. (Bottom) Summary of the alanine scanning mutagenesis of GAL4, which identify residues that are essential for activation by GAL11P in vivo. The levels of activation of each mutant were normalized against the wild-type LexA(1–202) + Gal4(50–97) derivative. Ala 91 in the wild-type GAL4 was replaced with a leucine. B] Control for the gel-shift experiments. GAL4(1–97) binds to its cognate site whereas neither GST–GAL11 nor GST–GAL11P binds this oligonucleotide. As shown in lanes 11–16, GST–GAL11P but not GST–GAL11 binds the GAL4–DNA complex, causing a further decrease in the electrophoretic mobility of the oligonucleotide–protein complex (supershift). GST–GAL11P263–352 was at the final concentration of 1.6, 3.2, and 4.8 µM in lanes 2–4 and 11–13, respectively and so was GST–GAL11P263–352 in lanes 5–7 and 14–16. GAL4 (30 nM) was present in lanes 9 and 11–16. GAL4 (15 nM) was added to lane 8, and 45 nM GAL4 was added to lane 10. (C–G) Each of the alanine substituted derivatives of GAL4 (1–97) binds to the cognate DNA-binding site, and their GST–GAL11P263–352 binding properties are approximated by the intensity of the supershift. In this experiment, 3 µM GST–GAL11P263–352 was added to each GAL4–DNA complex. GAL4 mutants were at ∼30 nM concentrations, at this concentration, each derivative saturably binds the labeled DNA probe.

**Figure 5.** Alanine scanning mutagenesis experiments identify GAL4 residues crucial for the activation in vivo and electrophoretic mobility shift assays identify GAL4 mutants that bind GAL11P263–352.
NMR structure of the GAL4 dimerization domain
necessary for transcriptional activation in GAL11P cells [Barberis et al. 1995]. To identify individual residues within this segment of GAL4 that play a role in transcriptional activation in vivo, we performed site-directed alanine substitution mutagenesis of each of the 40 residues [Fig. 5]. The mutations were made in the context of a chimeric protein in which residues 50–97 of GAL4 were fused to LexA, a dimeric bacterial DNA-binding protein that functions efficiently in yeast. In GAL11P cells, the LexA–GAL4 (50–97) fusion protein stimulates the expression of a reporter gene bearing LexA-binding sites upstream of the minimal GAL1 promoter [Barberis et al. 1995; Farrell et al. 1996]. Thus, it allowed us to circumvent any effects of alanine substitutions on dimerization and subsequent DNA binding of GAL4 to its cognate sites and identified those residues that specifically effect the ability of GAL4-dd to activate transcription. Figure 5A shows a scheme of the experimental design and a summary of the activity of each mutant normalized to the wild-type LexA–GAL4 (58–97) fusion protein. The results shown in Fig. 5A identify a cluster of hydrophobic residues from 69–72 (LLIF), and residues L64, E75, D76, S85, L86, I89, L92, L93, L96, and F97 as essential for activation in vivo.

To ascertain if the loss of activation occurs because of a loss in interaction, we purified wild-type GAL4 (1–97) and 39 mutant forms of each of which bears an alanine mutation at a single position scanning the residues 58–97, except for position 91, which is an alanine in the wild type and substituted by leucine. The mutant GAL4, D76A, was insoluble and thus refractory to further analysis. The in vitro interaction of these GAL4 derivatives with a fragment of GAL11 and GAL11P [residues 263–352] expressed as a GST fusion protein was measured in a gel mobility shift assay [Fig. 5B–G]. In this assay, neither GAL11 nor GAL11P bind to DNA bearing a GAL4 DNA-binding site [Fig. 5B, lanes 1–7], but, as expected, GAL4 (1–97) does [Fig. 5B, lanes 8–10]. Conversely, Figure 5B shows that GAL11P (lanes 14–16), but not GAL11 wild type [lanes 11–13], binds to the DNA-tethered GAL4 as indicated by the supershift of the protein–DNA complex. Figure 5(C–G) shows that all 39 GAL4 mutants bound the GAL4 DNA site, and 26 of these interacted with GAL11P as shown by the supershift.

There is a strong correlation between the ability of each GAL4 mutant to stimulate activation in vivo as measured in Figure 5A and the ability to interact with GAL11P in vitro as measured in Figure 5C–G. Thus, of the 39 mutants, those that interacted with GAL11P in vitro, all activated in vivo. Moreover, all of the 13 derivatives that failed to interact in vitro (positions 64, 69–72, 75, 85, 86, 89, 92, 93, 96, and 97) also failed to activate in vivo. K90A bound GAL11P very weakly in vitro but activated well in vivo. Conversely, E58A bound well in vitro but did not activate well in vivo; however, it was particularly difficult to purify, and thus its inability to activate transcription might be a consequence of instability in vivo.

The results from the alanine scanning are elucidated further by visualizing the crucial mutations [Fig. 7B, below] on the GAL4-dd structure [Fig. 1B]. Leu 69 to Phe 72 map on the hydrophobic part of the loop connecting α1 and α2 whereas E75 and D76 are on the hydrophilic section of the same loop. M83 to L86 constitute the loop between α2 and α3, and I89, L92, L93, L96, and F97 are in helix α3. Leu 64 form the core of α1 dimerization interface, and the effect of its mutation is probably purely structural.

**NMR titration of GAL4-dd with GAL11P261–352**

To identify the GAL11P-binding site, we titrated 15N-labeled GAL4-dd with unlabeled GAL11P261–352 and monitored the effect with 1H–15N HSQC spectra. Three of five spectra recorded in this series are shown in Figure 6B [0, 0.5, and 1 mM GAL11P]. Increasing concentrations of GAL11P from 0 to 2 mM, in 0.25-mM increments, were added to 1 mM [dimer] 15N-GAL4-dd while measuring 1H–15N HSQC spectra. Because only the signals of protons covalently attached to 15N nuclei are detected, all resonances observed arise from the 15N-labeled GAL4-dd while the GAL11P peptide is invisible. We obtained the following results.

GAL11P forms a tight complex with GAL4-dd with slow exchange kinetics (koff<5/sec) because adding GAL11P causes disappearance of a subset of the HSQC peaks and appearance of new peaks proportional to the concentration of the ligand [Fig. 6B]. Addition of the wild-type GAL11 peptide under otherwise identical conditions had no effect on the GAL4-dd resonances up to a 1:3 ratio of GAL4-dd to wt-GAL11 [data not shown].

GAL11P binds as a monomer to the GAL4-dd dimer because addition of GAL11P to the 1 mM [dimer] sample of GAL4-dd causes spectral changes up to 1 mM [monomer] GAL11P concentration. No significant spectral changes happen at higher concentration of GAL11P. The relative concentrations of the GAL4-dd versus GAL11P peptide particularly for this assays were verified by amino acid analysis.

Formation of the monomer–dimer complex introduces asymmetry into the GAL4-dd dimer [Fig. 6B]. Although the spectrum of the complex has not been assigned, a splitting of signals is observed for several residues, most clearly documented for G95, which is the only glycine and thus well separated from all other signals. The peak at 109.6 ppm 15N chemical shift splits in two peaks at 108.3 and 110.1 ppm 15N chemical shift. Many other peaks disappear in the titration and numerous new signals of half the intensity appear. Thus, introduction of asymmetry is manifested in many other peaks. However, the new peaks cannot easily be assigned because they are far away from those of the free GAL4-dd, and a complete assignment of the complex was not achieved. The assignments for those peaks that clearly disappear in the complex are highlighted in yellow in the left panel of Figure 6B, those that remain unchanged are not.
Figure 6. GAL11P binds the GAL4-dd dimer with slow exchange kinetics, and as a monomer. (A) Functional domains and partial amino acid sequence of GAL11P<sub>261-352</sub>. The N342V mutation is labeled in red. Position 342 is surrounded by a relatively high number of hydrophobic and basic residues. (B) NMR titration of <sup>15</sup>N-labeled GAL4-dd (50–106) with unlabeled GAL11P (261–352). <sup>1</sup>H, <sup>15</sup>N) HSQC spectra of 1.0 mM GAL4 [50–106] dimer alone [left], in the presence of 0.5 mM GAL11P [middle], and in the presence of 1.0 mM GAL11P [right]. The sample contained 50 mM phosphate buffer at pH 6.8, 100 mM NaCl. Cross peaks are labeled in the spectra of free GAL4 [left]. The signal of G95 that splits in two asymmetric peaks on GAL11P binding is marked with a box at the top of the spectra. Several other signals, primarily from helix α3 disappear in the presence of GAL11P. The assignment of these signals is highlighted in the left panel. Numerous new signals of lower intensity appear in the right panel manifesting widespread asymmetry of GAL4-dd in the complex with GAL11P.
Examination of the chemical shift perturbation of the NMR resonances on titration shows that there are no or only very minor changes observed for the entire helix α1. All residues of the loop connecting α1 and α2 disappear with increasing GAL11P concentration. The helix α2 is only affected at its C-terminal end by the addition of GAL11P. However, all of the helix α3 and the subsequent residues 97–100 disappear or split in two components. Because the binding of GAL11P peptide to GAL4-dd undoubtedly brings G95 to the vicinity of GAL11P, we assumed that the nearby residues whose resonances also are perturbed in the titration experiments also must be in close proximity to GAL11P. However, many of these GAL4 residues that participate in the dimerization may play a role in direct interactions with GAL11P, as observed in the case of G95. These observations indicate that GAL11P binds GAL4-dd via the α1–α2 loop and the helix α3.

Figure 7 shows the residues sensitive to GAL11P identified by NMR titration [Fig. 7A] and by alanine-scanning mutagenesis (Fig. 7B) mapped onto the molecular surface of GAL4-dd. Residues that are largely perturbed on GAL11P addition in the NMR titration are colored in red [Fig. 7A]. Two regions stand out: the hydrophobic α1–α2 loop and α3. Alanine substitutions that decrease activation in vivo to lower than 10% of the wild-type GAL4-dd are colored in yellow onto the same molecular surface representation of the GAL4-dd [Fig. 7B]. Comparison of A and B in Figure 7 reveals a remarkably close match between the red patches identified by NMR titration [left panel] with the yellow ones identified by alanine scanning [right panel]. Thus, in general the GAL4-dd residues shown to be sensitive to GAL11P in the NMR titration experiment are in close agreement with those shown to be crucial for transcriptional activation in vivo with GAL11P.

Discussion

GAL4-dd structure

The GAL4-dd structure is a C2-symmetric homodimer and contains three distinct coiled-coil helix pairs connected by mobile loops. The average structure is extended but mobile and poorly defined with respect to the twist angle around the C2 symmetry axis. Tilting motions that would bend the C2 symmetry axis are indicated by the distribution of the ensembles of calculated structures but must be fast on the NMR time scale because there is no solid evidence for asymmetry in the NMR spectra of the free GAL4-dd dimer. The roughly linear melting behavior, the absence of cooperative thermal unfolding, and the appearing and disappearing of HSQC peaks with variation of temperature indicate a unique dynamic nature of the GAL4-dd structure. This could be pictured as a transient melting and folding of the helix ends and a concomitant structural change of the connecting loops.

Complex structure

Our NMR titration data indicate that the dimeric GAL4-dd forms a tight complex with monomeric GAL11P261–352 with dissociation rates that are slow on the NMR time scale (k_{off} < 5/sec). This is consistent with a K_d of 300 nM or lower [Farrell et al. 1996; Simkovich 1996]. The monomer-to-dimer association renders the GAL4-dd as asymmetric, which is clearly documented by peak doubling in the NMR spectra [Fig. 6B]. This is a unique observation for the interaction of an activation domain with its target protein. Structural data on complexes between homodimers and monomers are rare but have been reported, such as the complex between human growth hormone and its dimeric receptor (de Vos et al.
or the RNA bacteriophage coat protein–operator complex [Valegard et al. 1994]. The symmetry-breaking aspects have been discussed extensively in Bardsley et al. [1998]. As discussed for the previous cases, binding of the first GAL11P monomer changes the conformation of the GAL4-dd dimer in a way that abrogates affinity for the second binding site. This is an extreme case of negative cooperativity.

Despite the tight interaction, the complex is in a dynamic state. This is concluded from the HSQC spectrum recorded at the end point of the titration [Fig. 6B, right] and from the spectra recorded with excess of GAL11P [data not shown]. These spectra contain far less than the number of signals expected for an entirely folded GAL4-dd dimer in complex with the unlabeled GAL11P. Conversely, an HSQC spectrum of free GAL11P shows only 48 weak peptide NH cross peaks [data not shown]. Thus, GAL4-dd, GAL11P, and their complex contain large portions that undergo apparently slow conformational exchange, which would make a more detailed structural characterization difficult.

**GAL4-dd shows a novel transcriptional activation motif**

The NMR titration and the alanine scanning experiments locate the GAL11P-binding site to the loop regions, the terminal helix α3, and part of the segment beyond α3. In wt-GAL4-dd, these regions bear six negative charges [E75, D76, D78, D84, D88, and D100] balanced by three basic side chains [K74, K82, and K90; Fig. 1B]. The double mutant [Q87R, K90E] that introduces an extra negative charge shows higher affinity to GAL11P and activates transcription stronger. Thus, the novel activation region of GAL4-dd shows the feature of high acidity present in other classic activating domains. Intriguingly, the GAL11P-binding face on GAL4-dd coincides with part of an acidic segment [residues 75-147] that previously been has shown to activate promoters bearing GAL4-binding sites in HeLa cell extracts [Lin et al. 1988]. Therefore, GAL11P may use the same activator motif.

In contrast with earlier studies of activating domains, which all revealed a single helix in the target-bound state [Kussie et al. 1996; Radakrishnan et al. 1997; Uesugi et al. 1997] the novel activation motif of the GAL4-dd is a coiled-coil dimer supporting two acidic/hydrophobic loops. The mobility experiments revealed that the GAL4-dd [the terminal α3 helix pair] shows enhanced mobility gradually increasing toward the C terminus [Fig. 4]. This may be important for activation function and is consistent with the observation that mutations stabilizing the α3 dimerization reduce the activation potential [A.Z. Ansari and M. Pashe, unpubl.].

The apparent need for weak dimer interaction of the α3 helices may explain why the double mutant version of the GAL4-dd studied here is a better activator. Both residues R87 and E90 are placed to the same side of the helix facing away from the dimerization interface. Interaction between the two side chains may stabilize the individual helices but not the dimer interaction. This is consistent with the previously mentioned observation that stabilizing the dimer reduces the activation function.

The recruitment of GAL11P by GAL4-dd relies on polar and hydrophobic interactions

The clustering of negative charges in the GAL11P-binding site of GAL4-dd suggest an electrostatic interaction. Consistently, the segment flanking the P-mutation in GAL11P has excess positive charge with five basic residues within ±10 positions [Fig. 6A]. It needs to be explained, however, why the N342V mutation in GAL11P generates a binding site for the GAL4-dd. The introduction of a hydrophobic residue at position 342 might cause folding of GAL11P, and the new binding function might be a conformational effect rather than a direct hydrophobic interaction of V342 with GAL4-dd. However, mutagenesis experiments in which residue at position 342 in GAL11 has been systematically substituted with various residues showed that only hydrophobic substitutions confer the transcriptional potentiator phenotype to GAL11. The different hydrophobic substitutions lead to different in vitro binding affinities for GAL4 and likewise different degrees of in vivo transcriptional activation [Barberis et al. 1995; Farrell 1996; Farrell et al. 1996]. It is difficult to imagine that such differences in affinities reflect different degree of structural stabilization toward permissive binding conformations. The large effect of such mutations at position 342 on binding affinity suggests a direct contact between the residue at this position with GAL4-dd. Thus, the structure of GAL4-dd and the features of the binding site suggest that the hydrophobic residue V342 and other nearby hydrophobic groups [Y340, I341, F336, I337, A445, L446, and I449; see Fig. 6A] in GAL11P might interact with the hydrophobic section of the loop connecting α1 and α2 in GAL4-dd. Mutational analysis indicates that in GAL11P bearing another hydrophobic substitution at position 342 [N342I], changing the neighboring hydrophobic residue I341 to Asn, reduces transcriptional activation some 10-fold. Moreover, adding the change Q343I to this GAL11 double mutant [N342I-I341N] improves activation by twofold. This observation further support the idea that hydrophobic residues neighboring 342 participate in GAL4–GAL11P interactions.

Hydrophobic interactions are usually tighter than hydrophilic ones. This might explain the low $K_d$ and the slow dissociation rates, which is unusual for activator–target interactions. For comparison, the VP16 activation domain is in fast exchange with the bound state, and the helical structure of the bound conformation was elucidated with transferred NOE experiment [Uesugi et al. 1997]. Typical activators are capable of contacting several targets to recruit the transcriptional machinery to the promoter [Pashne and Gann 1997; Koh et al. 1998 and references therein]. Thus, rapid exchange may allow them to recruit and orient the holoenzyme, and other components of the transcriptional machinery appropri-
GAL450–106 were analyzed by SDS-PAGE electrophoresis. 15N-(Amersham Pharmacia Biotech). The fractions containing column and gel filtration on a Superdex 75 (16/60) column followed by ion exchange chromatography on a DEAE–Sepharose ammonium sulfate precipitation with 0.25 and 0.15 g/mL, followed by centrifugation. Purification was performed through two consecutive steps at 37°C, induced 3 h with 1 mM IPTG, and harvested by centrifugation (Walters 1997a). The latter expresses poorly in bacteria. GAL450–106 was expressed from an rSETA expression vector (Invitrogen). Cells were grown at 37°C, induced 3 h with 1 mM IPTG, and harvested by centrifugation. The expression of GAL11P261–352 was performed similarly to the latter expresses poorly in bacteria. GAL450–106 was expressed in BL21-DE(3) pLysS Escherichia coli cells from a pSETA expression vector [Invitrogen]. Cells were grown at 37°C, induced 3 h with 1 mM IPTG, and harvested by centrifugation. Purification was performed through two consecutive ammonium sulfate precipitation with 0.25 and 0.15 g/mL, followed by ion exchange chromatography on a DEAE–Sepharose column and gel filtration on a Superdex 75 (16/60) column (Amersham Pharmacia Biotech). The fractions containing GAL450–106 were analyzed by SDS-PAGE electrophoresis. 15N- and 13C-labeled GAL4 were obtained by using 15NH4Cl and 15NH4Cl, 13C6-D-glucose as the sole nitrogen and carbon sources in M9 medium, respectively. H-15N-labeled protein was produced as described previously (Walters 1997a).

The expression of GAL11P261–352 was performed similarly from a pSETA vector in BL21-DE(3) pLysS E. coli cells. Purification of GAL11P261–352 was performed by partial fractionation with 0.3 g/mL ammonium sulfate followed by ion exchange chromatography on a SP-Sepharose and gel filtration on a Superdex-75 column.

Verification of the identity of the purified proteins was performed by mass spectrometry and amino acid analysis. The NMR protein samples ranged from 0.5 to 1 mM concentration in 50 mM sodium phosphate buffer at pH 7.4.

Alanine scanning mutagenesis

Mutations of the GAL4 segment in the context of LexA1–202 + Gal4(50–97) were generated by site-directed PCR mutagenesis. The PCR products were subcloned between XbaI–SalI sites in frame with the LexA coding sequence in RIR 238 bearing a chromosomal origin of replication, a centromere, and a His 3 histidine auxotrophy selection marker. The LexA derivatives are expressed by a strong yeast actin promoter and the transcriptional unit ends with a GAL11 terminator. Each of the 40 mutants was confirmed by sequencing. Expression was tested in a GAL11P strain from a chromosomally integrated β-gal reporter gene bearing two LexA-binding sites 50 bp upstream of the GAL1 promoter TATA box. β-gal assays were performed as described previously (Barberis et al. 1995; Smkovich 1996).

Electrophoretic mobility shift assay

A double stranded oligonucleotide of 30 bp containing a single GAL4-binding site (5′-TCCGAGAGCTGTCCCTCCGTT3′) was labeled with 32P at the 5′ end. The labeled DNA was incubated in a 20-μl reaction volume, with each of the GAL4 mutants and with GST–GAL11 or GST–GAL11P in EMSA buffer (200 mM Hepes at pH 7.5, 25 mM NaCl, 5 mM MgCl2, 20 mM ZnCl2, 6% glycerol, 200 μg/mL bovine serum albumin, 3 mM β-mercaptoethanol, 2 μg/mL polydI:dC). The complexes were resolved on a 9% polyacrylamide gel in 0.5× Tris-borate run for 2.5 h at 100 V at 9°C. The bands were visualized by Phosphor-imaging on a Fuji-Bas PhosphorImager.

The GAL4 mutants were expressed in BL21-DE(3) pLysS bacterial strain and purified on a SP-Fast Flow Sepharose resin (Pharmacia) as described previously (Ansari et al. 1998). GST–GAL11 or GST–GAL11P (residues 263–352 of GAL11) also were expressed in this strain and purified on a glutathione–Sepharose (Pharmacia) column by using standard procedures (Smith and Johnson 1988). Both proteins were dialyzed against the EMSA buffer.

Materials and methods

Sample preparation

Dr. R. Marmorstein kindly provided the purification procedure for GAL450–106. GAL450–106 was used instead of 50–97 because the latter expresses poorly in bacteria. GAL450–106 was expressed in BL21-DE(3) pLysS Escherichia coli cells from an rSETA expression vector [Invitrogen]. Cells were grown at 37°C, induced 3 h with 1 mM IPTG, and harvested by centrifugation. Purification was performed through two consecutive ammonium sulfate precipitation with 0.25 and 0.15 g/mL, followed by ion exchange chromatography on a DEAE–Sepharose column and gel filtration on a Superdex 75 (16/60) column (Amersham Pharmacia Biotech). The fractions containing GAL450–106 were analyzed by SDS-PAGE electrophoresis. 15N- and 13C-labeled GAL4 were obtained by using 15NH4Cl and 15NH4Cl, 13C6-D-glucose as the sole nitrogen and carbon sources in M9 medium, respectively. H-15N-labeled protein was produced as described previously (Walters 1997a).

The expression of GAL11P261–352 was performed similarly from a pSETA vector in BL21-DE(3) pLysS E. coli cells. Purification of GAL11P261–352 was performed by partial fractionation with 0.3 g/mL ammonium sulfate followed by ion exchange chromatography by using SP-Sepharose and gel filtration on a Superdex-75 column.

Identification of the purified protein was performed by mass spectrometry and amino acid analysis. The NMR protein samples ranged from 0.5 to 1 mM concentration in 50 mM sodium phosphate buffer at pH 7.4.

Alanine scanning mutagenesis

Mutations of the GAL4 segment in the context of LexA1–202 + Gal4(50–97) were generated by site-directed PCR mutagenesis. The PCR products were subcloned between XbaI–SalI sites in frame with the LexA coding sequence in RIR 238 bearing a chromosomal origin of replication, a centromere, and a His 3 histidine auxotrophy selection marker. The LexA derivatives are expressed by a strong yeast actin promoter and the transcriptional unit ends with a GAL11 terminator. Each of the 40 mutants was confirmed by sequencing. Expression was tested in a GAL11P strain from a chromosomally integrated β-gal reporter gene bearing two LexA-binding sites 50 bp upstream of the GAL1 promoter TATA box. β-gal assays were performed as described previously (Barberis et al. 1995; Smkovich 1996).

Electrophoretic mobility shift assay

A double stranded oligonucleotide of 30 bp containing a single GAL4-binding site (5′-TCCGAGAGCTGTCCCTCCGTT3′) was labeled with 32P at the 5′ end. The labeled DNA was incubated in a 20-μl reaction volume, with each of the GAL4 mutants and with GST–GAL11 or GST–GAL11P in EMSA buffer (200 mM Hepes at pH 7.5, 25 mM NaCl, 5 mM MgCl2, 20 mM ZnCl2, 6% glycerol, 200 μg/mL bovine serum albumin, 3 mM β-mercaptoethanol, 2 μg/mL polydI:dC). The complexes were resolved on a 9% polyacrylamide gel in 0.5× Tris-borate run for 2.5 h at 100 V at 9°C. The bands were visualized by Phosphor-imaging on a Fuji-Bas PhosphorImager.

The GAL4 mutants were expressed in BL21-DE(3) pLysS bacterial strain and purified on a SP-Fast Flow Sepharose resin (Pharmacia) as described previously (Ansari et al. 1998). GST–GAL11 or GST–GAL11P (residues 263–352 of GAL11) also were expressed in this strain and purified on a glutathione–Sepharose (Pharmacia) column by using standard procedures (Smith and Johnson 1988). Both proteins were dialyzed against the EMSA buffer.

NMR spectroscopy and assignments

All NMR spectra were recorded on Bruker AMX 500, Varian UnityPlus 400, Varian Unity 500, Inova 500, and UnityPlus 750 spectrometers at 35°C. Spectra were processed using FELIX (Molecular Simulations) and analyzed using XEASY (Bartels et al. 1995) on Silicon Graphics workstations. Backbone assignments were obtained using double-resonance 15N-dispersed NOE triple resonance three-dimensional 1H, 15N, and 13C HNCA and HN(CO)CA spectra. Side-chain resonances were assigned using 15N- and 13C-dispersed NOE experiments. HNHA and HNHB experiments were used to assign Hα and Hβ, respectively (Arch et al. 1991; Madsen et al. 1993; Vuiyster and Bax 1993). Intermonomer contacts were obtained as described previously (Walters et al. 1997a) using a 200-msec mixing time for the 15N-dispersed NOE. Complete mixing of the two homodimers species, unlabeled and 2H–15N-labeled, was achieved by heating the sample for 10 min to 60°C. No mixing was observed at room temperature over a period of 1 wk.

NOE and distance restraints

Distance restraints were obtained using a two-dimensional NOESY spectrum with a mixing time of 80 msec and 15N- and 13C-dispersed NOE (Talluri and Wagner 1996) spectra with mixing times of 120 and 80 msec, respectively. Cross peaks were integrated using the XEASY software package (Bartels et al. 1995). The intensities of NOE cross peaks arising from interactions between Hα and HN nuclei within helical regions were used to calibrate the 15N-dispersed NOE spectrum (Wüthrich 1986). The 13C-dispersed NOE spectrum was calibrated by assuming the weakest cross peak corresponding to an interaction between Hα and Hβ, within well-defined helical elements, to be 5 Å. NOE derived restraints were supplemented with 38 hydrogen bonds per monomer within regular secondary structural elements. Secondary structural elements were identified based on the Hα chemical shift index (Wishart et al. 1992) and NOE connectivity. For each hydrogen bond, two restraints were used such that the distances between the amide proton and the oxygen and between amide nitrogen and oxygen were restricted to be 1.8–2.5 Å and 2.5–3.3 Å, respectively. Distance symmetry restraints for all assigned residues were used to generated a twofold symmetric arrangement of the dimer as described previously (Brüger 1992).
Structure calculations

Structure calculations were performed using CNS version 0.5 [Brünger et al. 1998] on an R10000 Indigo II Silicon Graphics workstation. The molecular dynamics simulated annealing was performed using NMR distance restraints and distance symmetry restraints starting from an extended duplicated strand. With the exception of the NOESY cross peaks obtained using the GAL4 [50–106] heterodimer and the hydrogen bonds, all NOEs were treated as ambiguous [Nilges 1993]. The Insight II program [Molecular Simulations] and MOLMOL [Koradi et al. 1996] were used for displaying and visual inspection of the three-dimensional structures. PROCHECK (Laskowski et al. 1993) was used to further assess the quality of the structures. Residues in which the amide group was unassigned were dismissed from the statistics reported for the Ramachandran plot (Table 1). Coordinates of the structures have been deposited at the Protein Data Bank.

Relaxation experiments

All relaxation experiments were performed on a Varian Unity-Plus400 spectrometer. The three relaxation parameters, 15N longitudinal and transverse relaxation rates and heteronuclear NOE values, were measured using two-dimensional heteronuclear pulse sequences with pulsed-field gradients and sensitivity enhancement (Dayie and Wagner 1994). The analysis of the relaxation data was performed as described previously [Peng and Wagner 1992]. Each peak was integrated by volume by using XEASY and fitted to single exponential function by using the Levenburg–Marquardt nonlinear least squares method implemented in XEASY and fitted to a single exponential function by using the Levenburg–Marquardt nonlinear least squares method implemented in XEASY [New Unit]. Uncertainties in the relaxation rates were estimated with Monte Carlo simulations by using one repeated time point as described [Peng and Wagner 1992].

Acknowledgments

We are indebted to Dr. Kylie Walters for sharing her valuable experience in solving dimers structures throughout all this work. We are grateful to Dr. Yingxi Lin and Dr. Christoph Fahlie for many helpful discussions, to Gregory Heffron, Dr. Arne Persson, and Dr. Walfredo Antuch for their assistance in NMR spectroscopy. We thank the Helen Hay Whitney Foundation for support to A.Z.A. and Howard Hughes Medical Institute for a predoctoral fellowship to N.S. This work was supported by grants from the NIH and the NSF.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 USC section 1734 solely to indicate this fact.

References


Koradi, R., Billeter, M., and Wüthrich, K. 1996. MOLMOL: A NMR structure of the GAL4 dimerization domain


Recruitment of the transcriptional machinery through GAL11P: structure and interactions of the GAL4 dimerization domain

Patricia Hidalgo, Aseem Z. Ansari, Peter Schmidt, et al.

*Genes Dev.* 2001 15: 1007-1020

Access the most recent version at doi:10.1101/gad.873901

References

This article cites 46 articles, 11 of which can be accessed free at:
http://genesdev.cshlp.org/content/15/8/1007.full.html#ref-list-1

Email Alerting Service

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

To subscribe to *Genes & Development* go to:
http://genesdev.cshlp.org/subscriptions

Cold Spring Harbor Laboratory Press