MOLECULAR AND CELLULAR BIOLOGY, Sept. 1988, p. 3726–3733
0270-7306/88/093726-08$02.00/0

Vol. 8, No. 9

Cysteine Residues in the Zinc Finger and Amino Acids Adjacent to the Finger Are Necessary for DNA Binding by the LAC9 Regulatory Protein of *Kluyveromyces lactis*

MICHAEL M. WITTE AND ROBERT C. DICKSON*

Department of Biochemistry and Lucille Parker Markey Cancer Center, University of Kentucky, Lexington, Kentucky 40536-0084

Received 6 May 1988/Accepted 10 June 1988

LAC9 is a positive regulatory protein that controls transcription of the lactose-galactose regulon in *Kluyveromyces lactis*. LAC9 is homologous to the GAL4 protein of *Saccharomyces cerevisiae*. Both proteins have a single "zinc finger" which plays a role in DNA binding. We previously hypothesized (L. V. Wray, M. M. Witte, R. C. Dickson, and M. I. Riley, Mol. Cell. Biol. 7:1111–1121, 1987) that the DNA-binding domain of the LAC9 protein consisted of the zinc finger as well as a region of amino acids on the carboxyl-terminal side of the zinc finger. In this study we used oligonucleotide-directed mutagenesis to introduce 13 single-amino-acid changes into the proposed DNA-binding domain of the LAC9 protein. Variant LAC9 proteins carrying an amino acid substitution in any one of the four highly conserved Cys residues of the zinc finger had reduced DNA-binding activity, suggesting that each Cys is necessary for DNA binding. Three of four variant LAC9 proteins with amino acid substitutions located on the carboxyl-terminal side of the zinc finger had reduced DNA-binding activity. These results support our hypothesis that the DNA-binding domain of the LAC9 protein is composed of the zinc finger and the adjacent region on the carboxyl side of the zinc finger, a region that has the potential to form an α-helix. Finally, LAC9 proteins containing His residues substituted for the conserved Cys residues also had reduced DNA-binding activity, indicating that His residues are not equivalent to Cys residues, as had been previously thought.

DNA-binding proteins play critical roles in regulating gene expression. Biochemical, genetic, and structural studies of procaryotic regulatory proteins show that they use a common structural motif to bind DNA. The motif contains an α-helix followed by a turn and then a second α-helix (22). The specificity of binding is determined primarily by the first helix, the so-called recognition helix, which makes hydrogen bond and van der Waal contacts with bases in the major groove of the DNA molecule.

Some eucaryotic regulatory proteins use a helix-turn-helix motif to bind DNA. However, other DNA-binding motifs may exist. Recently, a structural motif for DNA-binding proteins, termed a metal or zinc DNA-binding finger, has been proposed from studies of the *Xenopus* transcription factor TFIIIA (21). TFIIIA contains nine repeats of an amino acid sequence with the consensus form of Cys-X_{2-5}-Cys-X_{12}-His-X_{2-3}-His. The protein also contains Zn^{2+} (6), and Miller et al. (21) proposed that the two Cys and the two His residues in each repeat sequence complex tetrahedrally to Zn^{2+} so that the intervening amino acids loop out to form a "finger" that contacts DNA. Since this initial report, many proteins have been shown to have one or more copies of a zinc finger with the general form Cys-X_{2-4}-Cys-X_{2-15}-*a*-X_{2-4}-*a* or *a*-X_{2-4}-*a*-X_{2-15}-Cys-X_{2-4}-Cys, where *a* is Cys or His (1). It remains to be determined how a single or multiple zinc fingers bind DNA.

We are examining how a *trans*-acting, positive regulatory protein, the LAC9 protein of the yeast *Kluyveromyces lactis*, regulates transcription of five structural genes which constitute the lactose-galactose regulon (23, 28). Genetic and biochemical evidence supports a model in which the LAC9 protein activates transcription by binding to a 17-base-pair (bp) upstream activator sequence (UAS) located in front of regulated genes (16, 25). It has been hypothesized that the transcriptional activator function of the LAC9 protein is modulated by the LAC10 protein. In the uninduced state the two proteins associate. In some unknown way inducers, lactose or galactose, cause the two proteins to dissociate, and the LAC9 protein then activates transcription (28).

The LAC9 protein, 865 amino acids, is homologous to the GAL4 protein, 881 amino acids, a positive regulatory protein of *Saccharomyces cerevisiae* (26, 28). The similarities and differences between these two proteins provide a unique opportunity to examine the function of eucaryotic regulatory proteins. The amino-terminal 74 amino acids of the GAL4 protein mediate binding to UAS (13), and 55% of the amino acids in this region are identical to those of the LAC9 protein. Two other regions, one of 174 amino acids near the middle of the proteins and one of 17 amino acids at their carboxy terminus, show 39 and 88% identity, respectively.

Besides structural similarities, the LAC9 and the GAL4 proteins have functional similarities. Expression of the lactose-galactose regulon of *K. lactis* (24) and the melibiose-galactose regulon of *S. cerevisiae* (26, 28) can be activated by either *LAC9* or *GAL4*. However, the proteins have unique properties and do not simply mimic each other. For example, the melibiose-galactose regulon is severely repressed by glucose, but this repression can be partially released if *LAC9* is substituted for *GAL4* (26, 28). The lactose-galactose regulon is normally not severely repressed by glucose but becomes so when *GAL4* is substituted for *LAC9* (24). We hypothesized that the GAL4 protein has a domain that responds to the glucose repression circuit, while either the LAC9 protein lacks such a domain or the domain is shielded (24, 28). Another difference between the two proteins is that the GAL80 protein modulates the transcrip-

* Corresponding author.

TABLE 1. Oligonucleotides for introducing base changes into *LAC9*

Oligonucleotide length (mer)	Nucleotide(s) mutated[a]	Amino acid substitution
23	+284, 285	Cys-95(TGC)→Ser(TCG)
25	+293, 294	Cys-98(TGC)→Ser(TCG)
23	+293	Cys-98(TGC)→Phe(TTC)
23	+292, 293	Cys-98(TGC)→Val(GTC)
23	+292, 293	Cys-98(TGC)→His(CAC)
22	+334	Cys-112(TGC)→Gly(GGC)
22	+334, 335	Cys-112(TGC)→His(CAC)
30	+344, 345	Cys-115(TGT)→Ser(TCG)
30	+343, 344	Cys-115(TGT)→His(CAT)
25	+374	Arg-135(AGA)→Thr(ACA)
25	+377	Ala-136(GCA)→Gly(GGA)
30	+388	Glu-140(GAG)→Lys(AAG)
30	+394	Glu-142(GAA)→Lys(AAA)

[a] Nucleotide numbers are relative to A in the ATG initiation codon.

tional activator activity of the GAL4 protein but does not seem to do so for the LAC9 protein (28).

We previously hypothesized that the DNA-binding domain of the GAL4 and the LAC9 proteins was composed of both the zinc finger and a region of the proteins located on the carboxyl side of the zinc finger (28). This hypothesis was based on an amino acid comparison of the LAC9 and the GAL4 proteins to the yeast regulatory proteins PPR1 (12) and ADR1 (7). Since these proteins have similar amino acid sequences in their zinc fingers, we reasoned that DNA-binding specificity must lie elsewhere in the proteins. Furthermore, we reasoned that the amino acids in the LAC9 and the GAL4 proteins that make specific contacts with DNA should be nearly identical, since the proteins contact the same UAS (16, 24, 28). A search for identical amino acids outside the zinc finger but within the first 74 amino acids of GAL4, the region known to mediate DNA binding, revealed a region located between 18 and 29 amino acids on the carboxy side of the zinc finger in which 10 of 12 amino acids are identical between GAL4 and LAC9. The corresponding region in PPR1 and ADR1 showed no similarity, as would be expected because these proteins bind different UAS. Another appealing attribute of this region of 12 amino acids was that it was predicted (28) to form an α-helix and could thus make hydrogen bond contacts in the major groove of DNA, just like the recognition helix of DNA-binding proteins that contain a helix-turn-helix motif.

To gain a better understanding of how proteins with a single zinc finger recognize a DNA sequence, we changed specific amino acid residues in the LAC9 protein and measured DNA-binding activity. Our results suggest that the DNA-binding domain of the LAC9 protein contains both the zinc finger and the potential α-helical region located 18 to 29 amino acid residues on the carboxyl side of the zinc finger.

MATERIALS AND METHODS

Oligonucleotide-directed mutagenesis. Mutagenic oligonucleotides (Table 1) were synthesized on an Applied Biosystems 380B automated DNA synthesizer. The template for mutagenesis, M13mp19-LAC9, was constructed by inserting a 3.6-kilobase (kb) *Eco*RI fragment carrying *LAC9* (subclone ΔB [28]) into the *Eco*RI site of M13mp19, oriented so that the transcribed strand of *LAC9* was packaged into phage. Oligonucleotide-directed mutagenesis (13a) was used to introduce specific missense mutations into the proposed DNA-binding domain of the *LAC9* gene. Briefly, M13mp19-LAC9 single-stranded DNA was prepared from phage grown on *Escherichia coli* BD2399 (*dut ung*; Bruce Duncan, Institute for Cancer Research, Philadelphia, Pa.). The uracil-containing phage DNA produced in this host was used as template for oligonucleotide-primed DNA synthesis. The resulting double-stranded DNA was transformed into the *dut*$^+$ *ung*$^+$ *E. coli* MV1190 {Δ(*lac-proAB*) *thi-1 supE* Δ(*srl-recA*)*306*::Tn*10* (Tet^r)[F′ *traD36 proAB lacI*^{q}*ZΔM15*]} to select against the uracil-containing template strand. The frequency of mutants ranged from 30 to 80%. The sequence of each mutation was confirmed by dideoxy sequencing (27).

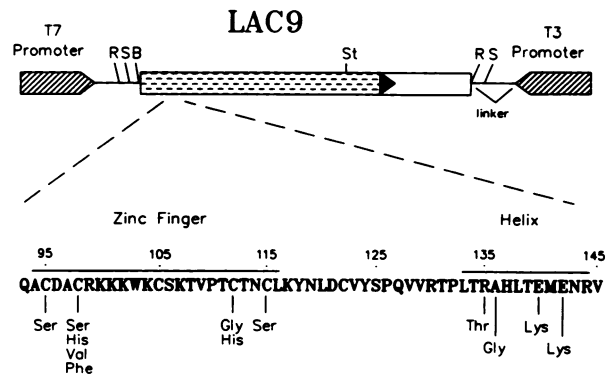

FIG. 1. Diagrammatic representation of *LAC9* and a portion of its coding region. Orientation of *LAC9* relative to the T7 promoter in the plasmid pBS-LAC9 (top) and an expanded view of the LAC9 protein sequence containing the zinc finger and potential α-helix region (bottom). Numbers above the amino acid sequence indicate amino acid residues relative to the amino-terminal methionine residue. Amino acid changes resulting from oligonucleotide-directed mutagenesis are shown below the amino acid sequence. The drawing is not to scale. Abbreviations for restriction endonuclease sites are as follows: R, *Eco*RI; S, *Sac*I; B, *Bgl*II; St, *Sty*I. Linker refers to the polylinker region of the Bluescribe vector (pBS) containing the restriction endonuclease sites *Kpn*I, *Xma*I, *Sma*I, *Bam*HI, *Xba*I, *Hin*cII, *Sal*I, *Acc*I, *Pst*I, *Sph*I, and *Hin*dIII.

Plasmid construction. The vector pBS-LAC9 (Fig. 1) was constructed by inserting a 3.4-kb *Bql*II-*Eco*RI fragment, carrying a promoterless *LAC9* gene, from p1B3-LAC9RΔB (28) into the *Bgl*II-*Eco*RI sites of pIC20H (18) to generate pIC20H-LAC9. A 3.4-kb *Sac*I fragment from pIC20H-LAC9 was inserted into the *Sac*I site of pBS M13$^+$ (Stratagene Cloning Systems, San Diego, Calif.) so that the *LAC9* coding sequence was correctly oriented and just downstream from the T7 phage promoter. Mutant *LAC9* sequences were inserted into pBS-LAC9 by replacing the wild-type 2.2-kb *BGl*II-*Sty*I fragment (Fig. 1) with the analogous fragment obtained from a M13mp19-LAC9 mutant sequence. The region around the mutation in each pBS-LAC9 construction was confirmed by dideoxy sequencing of double-stranded plasmid DNA (3). To avoid the possibility of base changes occurring outside the sequenced region during the numerous in vitro and in vivo manipulations, two or three independent isolates of each mutation, except Glu-140 and Cys-115, were isolated in M13, introduced into pBS, and tested for their effects on DNA binding. The double mutation Cys-112→His plus Cys-115→His was constructed by mutating first Cys-112 and then Cys-115.

Multicopy *S. cerevisiae* expression vectors carrying wild-type or mutant *LAC9* coding sequences under control of the *ADH1* promoter were constructed by inserting a 3.4-kb *Bgl*II-*Hin*dIII fragment, from either wild-type or mutant

pBS-LAC9, into the *Bam*HI-*Hin*dIII sites of p1AJ(S) (Oh and Hopper, University of Pennsylvania, Hershey, Pa.). Plasmid 1AJ(S) was constructed by inserting a *Sph*I-*Hin*dIII DNA fragment carrying the *GAL4* coding region fused to the *ADH1* promoter into the *Sph*I-*Hin*dIII sites of pJDB207. Plasmid 1AJ(S) has a *LEU2* gene for selection of yeast transformants and a 2μm origin of replication.

Single-copy *S. cerevisiae* vectors carrying wild-type or mutant *LAC9* coding sequences under the control of the normal *LAC9* promoter were constructed in two steps. First, mutant *LAC9* coding sequences were joined to the normal *LAC9* promoter by replacing the wild-type 729-bp *Bgl*II-*Kpn*I fragment of pRS2 (28) with the analogous fragment from a pBS-*LAC9* mutant sequence. Second, a 5.4-kb *Eco*RI fragment from the pRS2 construct, carrying either wild-type or mutant *LAC9* sequences, was inserted into the *Eco*RI site of the single-copy vector YCp50 (28) to generate YCp50-*LAC9*. Plasmids were introduced into *S. cerevisiae* SJ21R2 by a lithium acetate transformation procedure (28). SJ21R2 (**a** *gal4::ura3 leu2-3,112 ade1 ura3-52 MEL1*) is a Ura$^-$ derivative of a strain in which the region between the *Sna*BI (51 bp in front of the ATG initiation codon) and the *Nru*I (at bp 999) restriction sites in the *GAL4* gene was deleted and replaced with the *URA3* gene by transplacement. The Gal phenotype was determined on minimal galactose plates (23).

In vitro transcription and translation. Mutant and wild-type pBS-LAC9 were linearized with the restriction enzyme *Bam*HI, which cleaves in the linker at the 3′ end of *LAC9* (Fig. 1). DNA template (2½ μg) was transcribed in vitro by using T7 RNA polymerase in a 25-μl reaction as described by Melton et al. (19). The total RNA generated in such a reaction (roughly 2.5 μg) was treated with DNase I (Organon Teknika, Malvern, Pa.; LS06333, RNase free, 1 U/μg of DNA), extracted with phenol-chloroform and chloroform, and ethanol precipitated. RNA was translated by using a rabbit reticulocyte lysate in a 10-μl reaction containing 10 μCi of [^{35}S]methionine (1,129 Ci/mmol; New England Nuclear Corp., Boston, Mass.) as described by Hellmann et al. (8) or as described by Promega Biotec (1987, Madison, Wis.).

Specific synthesis of LAC9 protein was obtained. Starting with approximately 125 ng of the synthetic LAC9 mRNA, typically 3×10^5 cpm of [^{35}S]methionine was incorporated into LAC9 protein in a translation volume of 10 μl. From the predicted amino acid sequence of the LAC9 protein (28) and the specific activity of the [^{35}S]methionine, a minimum of approximately 5 ng of full-length LAC9 protein was synthesized per reaction. This corresponds to a LAC9 protein concentration of approximately 5 nM. Inclusion of the cap analog diguanosine triphosphate (G-5′ppp5′-G, 500 μM) in the transcription reaction (9) did not increase the amount of full-length LAC9 protein synthesized. Consequently, G-5′ppp5′-G was not routinely used.

Translation products were examined by electrophoresis in a 12.5% polyacrylamide gel (acrylamide: bisacrylamide, 30: 0.8) containing sodium dodecyl sulfate as described by Laemmli (14). Low-molecular-weight protein standards (10 to 100 kilodaltons; Bio-Rad Laboratories, Richmond, Calif.) were used as markers. After electrophoresis, the gel was fixed in 10% trichloroacetic acid–10% acetic acid for 1 h. Proteins were stained with 0.1% Coomassie brilliant blue R in 50% trichloroacetic acid for 30 min, and the gel was destained in 7% acetic acid overnight. The gel was treated with 100% dimethyl sulfoxide three times for 20 min each, impregnated for 1 h with 1 M 2,5-diphenyloxazole (PPO) in dimethyl sulfoxide, washed with water for 1 h, dried, and fluorographed.

DNA-binding assay. DNAs were purified by equilibrium centrifugation in CsCl. Binding experiments used pIC20R-UAS DNA, which contains a 22-bp synthetic oligonucleotide carrying a 17-bp UAS from the *LAC4* promoter (16) cloned into the *Sal*I site of pIC20R (18). Both pIC20R-UAS and pIC20R were digested with *Pvu*II, and the 344- or 322-bp fragment, respectively, was isolated by electroelution from a 2% agarose gel.

The DNA-binding assay was similar to that described by Hope and Struhl (9) except that the final DNA concentration was 20 nM. After electrophoresis, the gel was fixed, treated with dimethyl sulfoxide, impregnated with PPO, dried, and fluorographed.

RESULTS

DNA-binding activity of LAC9. Using oligonucleotide-directed mutagenesis of *LAC9* (13a), we introduced 13 single-amino-acid changes into the proposed DNA-binding domain of the LAC9 protein. Sequence-specific binding of the variant and wild-type LAC9 proteins was determined by in vitro techniques similar to those described by Hope and Struhl (9). These techniques were used because they produce enough pure protein to measure DNA-binding activity, and many variant proteins containing amino acid changes can be made and assayed in a short time. We produced LAC9 proteins in the following manner. A hybrid DNA molecule consisting of the wild-type or mutated *LAC9* coding region cloned into the pBS vector (Fig. 1) was cleaved with *Bam*HI, and the linear DNA was transcribed with purified T7 RNA polymerase to produce a runoff transcript. The runoff transcript was translated in vitro by using a rabbit reticulocyte lysate system to generate [^{35}S]methionine-labeled LAC9 protein.

Translation products for the wild-type LAC9 protein and each variant LAC9 protein are shown in Fig. 2. In every case, a translation product of 97 kilodaltons was seen. This is the predicted molecular mass for the LAC9 protein (28). These results demonstrate that the amino acid changes do not prevent the synthesis of full-length LAC9 protein. Furthermore, the sodium dodecyl sulfate gel electrophoresis results enabled us to assay equal amounts of full-length wild-type and variant proteins. Bands of lower molecular weight than that of full-length LAC9 were also observed. They probably correspond to proteins arising from initiation of translation at internal methionine codons, since they are the exact size predicted for such products. The presence of these protein fragments did not interfere with the DNA-binding assay (see below).

The DNA-binding activity of wild-type and variant LAC9 proteins was measured by using a band shift assay (9). The DNA substrate was a 344-bp fragment carrying a 17-bp UAS. DNA was incubated with radioactive LAC9 protein, and the sample was subjected to electrophoresis in a polyacrylamide gel. Under the conditions used for this assay, the full-length radioactive LAC9 protein should, because of its net neutral charge, enter the gel matrix only if it has bound to the DNA. The position of the protein-DNA complex on the gel was revealed by fluorography. In the presence of the UAS-containing DNA fragment, a strong band of ^{35}S-labeled LAC9 protein migrated into the gel (Fig. 3, lane 5). This band, indicated by an arrow in Fig. 3, represents LAC9 protein bound to its UAS, since the band was not seen if DNA was omitted (Fig. 3, lane 4) or if the same DNA

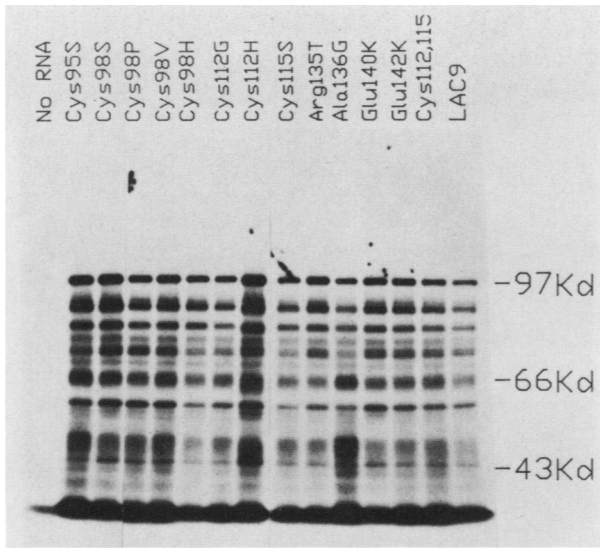

FIG. 2. Synthesis of LAC9 proteins. Autoradiograph of in vitro translation products analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. RNA for translation reactions was prepared by in vitro transcription of wild-type and mutant alleles carried in pBS-LAC9. Equal volumes of the translation products per lane were loaded. No exogenous RNA was added to the translation reaction marked No RNA. Size markers (kilodaltons [Kd]) were determined by comparison to proteins of known molecular weights (not shown).

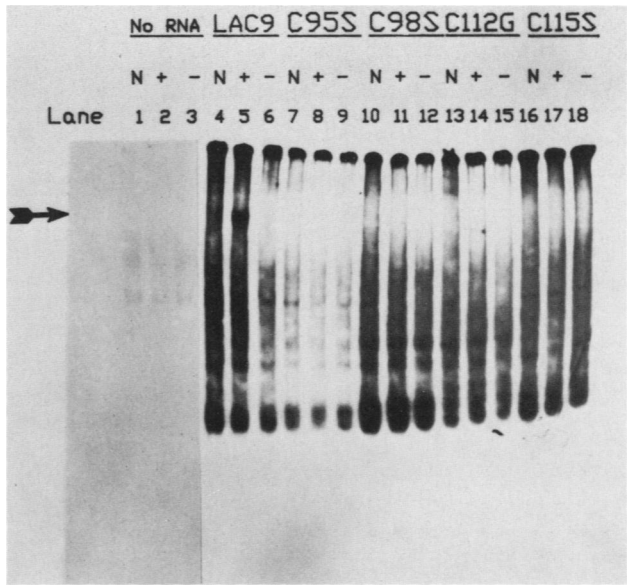

FIG. 3. Cysteine residues in the zinc finger essential for DNA binding. ^{35}S-LAC9 proteins were generated by transcription and translation of wild-type or mutated pBS-LAC9 DNAs. Translation products produced in the absence of LAC9 mRNA (lanes 1 through 3) or from wild-type LAC9 mRNA (lanes 4 through 6) or mutated LAC9 mRNAs carrying substitutions at Cys-95 (lanes 7 through 9), Cys-98 (lanes 10 through 12), Cys-112 (lanes 13 through 15), or Cys-115 (lanes 16 through 18) were each incubated either with a 344-bp DNA fragment carrying a UAS (+), with the same DNA fragment (322 bp) lacking a UAS (−), or with no added DNA (N). Protein-DNA complex formation, indicated by the arrow, was analyzed by native polyacrylamide gel electrophoresis and autoradiography.

fragment lacking the UAS was used (Fig. 3, lane 6). Also, the ^{35}S-LAC9 protein band was not seen when the translation reaction was performed in the absence of LAC9 RNA (Fig. 3, lanes 1 through 3), indicating that the radioactive protein-DNA complex was dependent upon LAC9 mRNA and was not due to a component in the rabbit reticulocyte lysate. These results show that the LAC9 protein binds specifically to DNA carrying a UAS.

Cysteine residues in the zinc finger essential for DNA binding. We first wanted to demonstrate directly that the zinc finger of the LAC9 protein was needed for DNA binding. Since it has not been shown that each of the four conserved Cys in the zinc finger of a protein with one finger are essential for DNA binding as proposed in the zinc finger model of Miller et al. (21), we changed each Cys, one at a time, and measured DNA binding of the variant proteins. Our rationale for choosing amino acids to replace Cys residues was to prevent coordination to Zn^{2+} yet retain the normal conformation of the DNA binding domain. Consequently, we replaced Cys with the small amino acid Ser or Gly; larger and chemically different residues were also introduced to try to disrupt the conformation of the domain and thus reduce DNA binding (Fig. 1). In each case, substitution of Cys-95, Cys-98, Cys-112, or Cys-115 with one of the amino acids shown in Fig. 1 reduced DNA-binding activity (Fig. 3). We conclude that the zinc finger is part of the DNA-binding domain of the LAC9 protein and that each Cys residue of the proposed zinc finger is essential for normal DNA-binding activity.

Effects of amino acids outside the zinc finger on DNA binding. We hypothesized (28) that amino acids located 18 to 27 residues on the carboxyl side of the zinc finger are part of the LAC9 DNA binding domain. This region of the protein will be referred to as the 18-27 or potential α-helix-forming region (Fig. 1). To test this hypothesis, we constructed variant LAC9 proteins with a single-amino-acid substitution in the 18-27 region. The LAC9 variants carrying the amino acid substitution Arg-135→Thr, Ala-136→Gly, or Glu-142→Lys did not bind in vitro to DNA containing a UAS (Fig. 4A). These results suggest that residues 135, 136, and 142 of LAC9 are part of the DNA-binding domain. The LAC9 variant carrying the substitution Glu-140→Lys maintained DNA-binding activity (Fig. 4A, lane 14), indicating that Glu-140 does not make specific H-bond or van der Waals contacts to DNA. In these experiments, multiple protein-DNA complexes were visible in both the wild-type LAC9 and the Glu-140→Lys protein samples. The nature of the faster migrating band is unclear, but it must involve the LAC9 protein and UAS, since it appears only when the normal, slower migrating complex is present. For reasons we do not understand, the presence of the faster migrating band was dependent upon the batch of rabbit reticulocyte lysate used for translation.

To determine if the loss of detectable DNA binding was due simply to a small reduction in binding affinity of the variant proteins for UAS, we assayed protein binding at the normal DNA concentration of 20 nM and at a 5- or 10-fold-higher concentration. At these higher concentrations the Ala-136→Gly variant LAC9 protein showed DNA-binding activity (Fig. 4B), suggesting that the amino acid substitution reduced DNA binding about fivefold. The Arg-135→Thr and the Glu-142→Lys variant proteins did not show binding activity even at the 10-fold-higher DNA concentration, indicating at least a 10-fold reduction in DNA-binding activity.

The DNA-binding activity of these variant proteins was

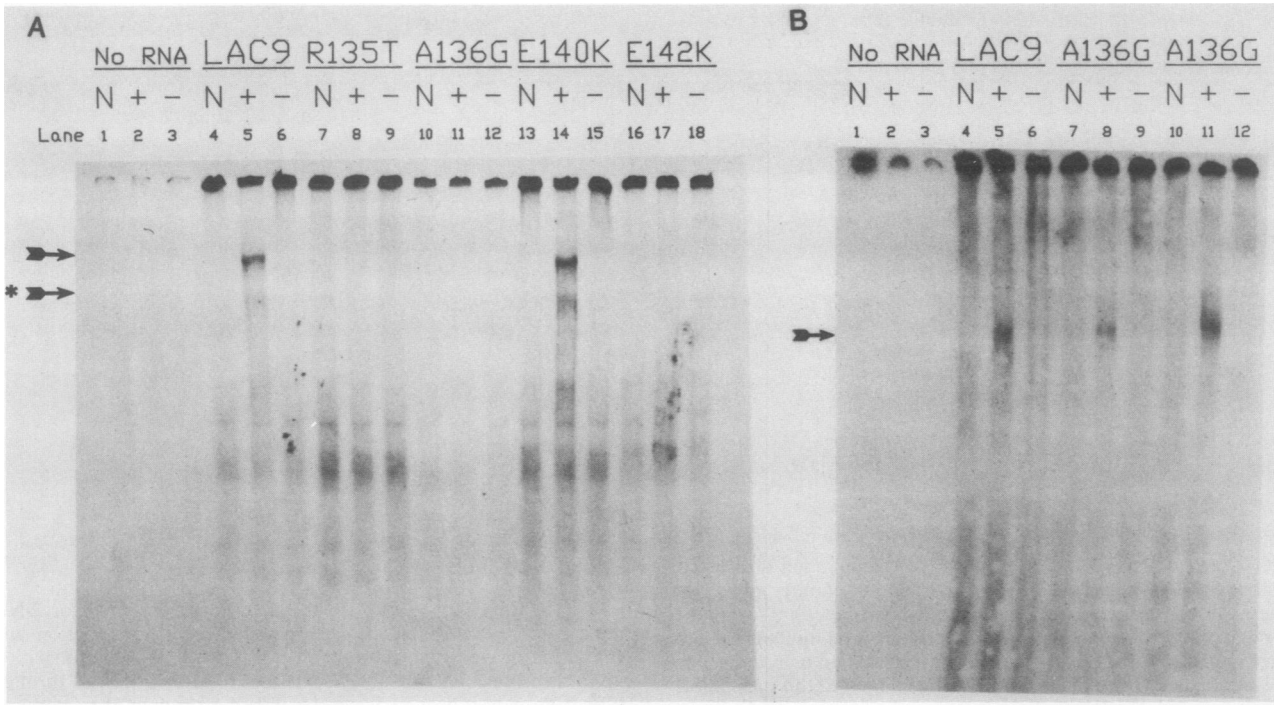

FIG. 4. Effects of amino acids outside the zinc finger on DNA binding. (A) ^{35}S-labeled translation products produced by in vitro translation of wild-type LAC9 mRNA (lanes 4 through 6) or variant LAC9 mRNAs carrying substitutions at Arg-135 (lanes 7 through 9), Ala-136 (lanes 10 through 12), Glu-140 (lanes 13 through 15), or Glu-142 (lanes 16 through 18) or in the absence of LAC9 mRNA (lanes 1 through 3) were each incubated with DNA (20 nM) carrying a UAS (+), with DNA without a UAS (−), or with no DNA (N). Protein-DNA complex formation, indicated by the arrow, was analyzed by native polyacrylamide gel electrophoresis and autoradiography. The faster-migrating protein-DNA complex is indicated by the arrow with an asterisk. (B) ^{35}S-labeled Ala-136 variant LAC9 protein was assayed for reduced binding affinity for UAS by using the DNA binding assay as described above, except that the final DNA concentrations of the fragments to be bound were 5-fold (100 nM [lanes 7 through 9]) or 10-fold (200 nM [lanes 10 through 12]) higher than those used with the wild-type LAC9 protein (20 nM [lanes 4 through 6]). Protein-DNA complex formation is indicated by the arrow.

also examined in vivo by determining whether the variant protein could activate transcription of the galactose regulon in a *gal4* deletion strain of *S. cerevisiae* and allow growth of the strain on minimal galactose plates. *S. cerevisiae* was used as a host because both single-copy and multicopy vectors are available, thus allowing us to vary the amount of protein per cell by varying the number of copies of the gene. Only multicopy vectors are available for *K. lactis*. Also, by using *S. cerevisiae* we could overproduce the variant proteins by using the highly expressed *ADH1* promoter. Each of the four mutant *LAC9* genes carrying the normal *LAC9* promoter was inserted into the single-copy vector YCp50. To try to obtain increased concentrations of each variant protein, the coding region of the mutant *LAC9* genes was inserted into a multicopy vector in front of the *ADH1* promoter. The vectors were transformed into the *gal4* deletion strain SJ21R2 and tested for growth on minimal galactose medium by replica plating. Wild-type *LAC9* and the Glu-140→Lys variant gave Gal$^+$ cells when present on a single- or a multicopy vector, indicating that the *LAC9* promoter was expressed well enough by *S. cerevisiae* to give strong activation of the galactose regulon even when only one copy of *LAC9* per cell was present. The Ala-136→Gly variant gave very slow growth when present on the single-copy vector but gave wild-type growth when the mutant gene was fused to the *ADH1* promoter and carried on a multicopy vector. In view of the in vitro DNA-binding data, we believe that the Gal$^+$ phenotype is due to overproduction of the Ala-136→Gly variant protein and consequent activation of the galactose regulon. The Arg-135→Thr and the Glu-142→Lys variants did not give Gal$^+$ cells even when the mutant gene was fused to the *ADH1* promoter and placed on a multicopy vector. These results are in agreement with the lack of binding shown by the in vitro DNA-binding assay (Fig. 4).

Johnston (10) showed that the Gal$^-$ phenotype of a *gal4* mutant strain carrying a Pro-26→Leu change in the *GAL4* protein could be converted to Gal$^+$ by high concentrations of $ZnCl_2$ in the culture medium. Presumably the reduced DNA-binding activity of the Pro-26→Leu variant protein was restored by a high concentration of Zn^{2+}. To determine whether the DNA-binding activity of the Arg-135→Thr or the Glu-142→Lys variant LAC9 proteins could be restored by a high Zn^{2+} concentration, we transformed strain SJ21R2 with the multicopy vector carrying the *LAC9* mutant coding region fused to the *ADH1* promoter. Transformed cells were examined for growth on minimal galactose plates containing a crystal of $ZnCl_2$ at the edge of the plate. Under these conditions the Zn^{2+} gradient did not restore the Gal$^+$ phenotype to strains producing either variant protein.

Inability of histidine to substitute for cysteine in the zinc finger of LAC9. A number of proteins have zinc fingers in which one or two His residues are found instead of Cys; most often, the second pair of Cys residues is replaced by His residues (1). To determine whether His can replace Cys in the zinc finger of the LAC9 protein, we constructed *LAC9* mutants with His in place of the second (Cys-98), third (Cys-112), or third and fourth (Cys-115) Cys residues of the

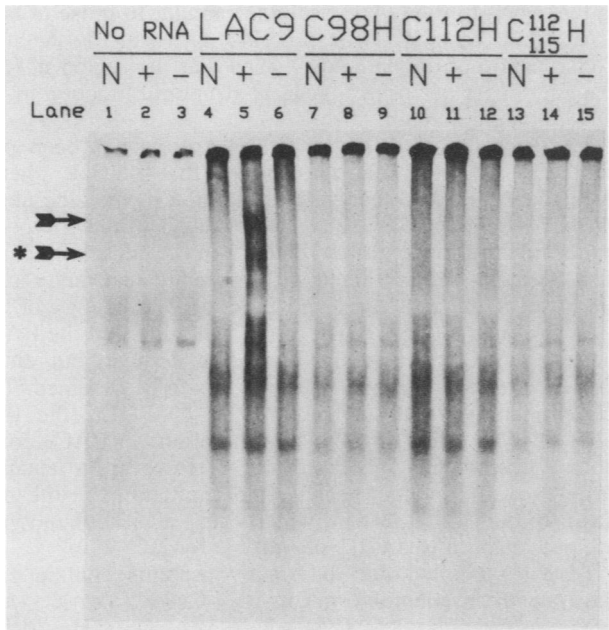

FIG. 5. Inability of histidine to substitute for cysteine in the zinc finger of LAC9. ^{35}S-labeled wild-type (lanes 4 through 6) and variant LAC9 proteins carrying His substituted for Cys-98 (lanes 7 through 9), Cys-112 (lanes 10 through 12), or both Cys-112 and Cys-115 (lanes 13 through 15) produced by in vitro transcription and translation of pBS-LAC9 DNAs or in the absence of LAC9 mRNA (lanes 1 through 3) were incubated with DNA carrying a UAS (+), with DNA without a UAS (−), or no DNA (N). Protein-DNA complex formation, indicated by the arrow, was analyzed by native polyacrylamide gel electrophoresis and autoradiography. The faster-migrating protein-DNA complex is indicated by the arrow with an asterisk.

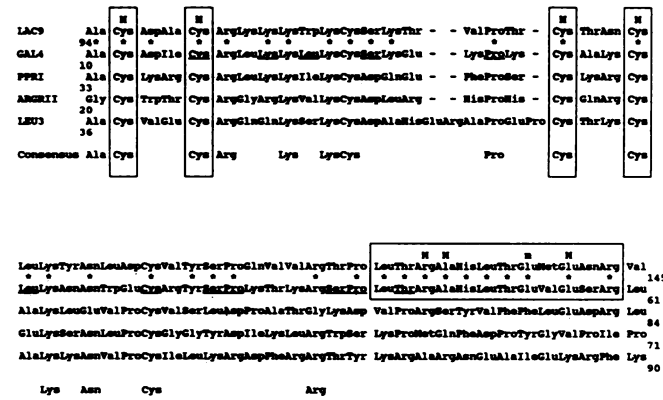

FIG. 6. Yeast regulatory proteins with one zinc finger. Amino acid sequence of the zinc finger and adjacent region of the yeast regulatory proteins LAC9 (26, 28), GAL4 (15), ARGRII (20), and LEU3 (29) are compared. The four invariant Cys residues that are proposed to contact Zn^{2+} are indicated by vertical boxes. The potential α-helical region of LAC9 and GAL4 is indicated by the horizontal box. Symbols: *, amino acid identities between LAC9 and GAL4; M, amino acid changes (Fig. 1) that reduced the DNA-binding activity of the LAC9 protein; m, a change that did not affect DNA-binding activity. Changes in the underlined residues of GAL4 reduce DNA binding (11).

zinc finger. None of the variant LAC9 proteins bound to UAS (Fig. 5). These results show that His and Cys are not equivalent residues in the zinc finger of the LAC9 protein. Again, the wild-type LAC9 protein appeared to form two major protein-DNA complexes (Fig. 5, lane 5) because of the batch of rabbit reticulocyte lysate used for translation reactions. Each of these variant proteins was assayed at both 20 and 100 nM DNA to test for reduced affinity for UAS. No detectable DNA binding was observed with any variant LAC9 protein (data not shown), suggesting that their affinity for DNA had been reduced by more than fivefold.

DISCUSSION

Our results demonstrate that the DNA-binding activity of the LAC9 protein is reduced when any of the four Cys residues in the zinc finger is replaced by another small, uncharged amino acid, e.g., Ser or Gly. These results support the hypothesis of Miller et al. (21), who proposed that four Cys, or a combination of Cys and His residues, complex to Zn^{2+} to form a DNA-binding motif. Other data support this model. Johnston and Dover (11) showed that mutation of Cys-14→Tyr in the zinc finger of the GAL4 protein (Fig. 6) reduced DNA-binding activity. Blumberg et al. (2) found that the mutations Cys-106→Tyr, Cys-109→Tyr, His-118→Tyr, or His-122→Tyr in the first zinc finger and the mutation Cys-134→Tyr in the second zinc finger of the *S. cerevisiae* ADR1 protein reduced the ability of the protein to activate expression of a target gene. Presumably the variant ADR1 proteins failed to bind to UAS in front of the target gene. Taken together these data provide strong genetic and biochemical evidence for the central role of Cys or His residues in the DNA-binding function of zinc fingers.

A fundamental question about proteins with one zinc finger is that of the role of the zinc finger in DNA binding. Does the finger contact specific bases, or does it perform a more general role, such as displacing histones from DNA and contacting phosphates in the backbone of the DNA? If base-specific binding is determined by amino acids in the zinc finger, then residues between the two internal Cys residues, as proposed by Miller et al. (21), should be unique. This may not be the case for yeast DNA-binding proteins with one zinc finger. There is considerable similarity of the amino acids between the internal Cys residues of the zinc finger in the LAC9, GAL4, PPR1, ARGRII, and LEU3 proteins (there is no direct evidence that the indicated region of PPRI, ARGRII, or LEU3 is the DNA-binding domain) (Fig. 6). Since the LAC9 and GAL4 proteins probably bind to the same UAS (24, 26, 28), we would expect the amino acids in them that make contacts with DNA to be identical or nearly so. These amino acids should differ from those in PPR1, ARGRII, and LEU3, since these three proteins activate expression of unique target genes probably by recognizing unique UAS. Figure 6 shows a sequence of 12 amino acids in LAC9 and GAL4 located 18 to 27 residues on the carboxyl side of the zinc finger in which 10 of the residues are identical. To determine whether the 18-27 region of the protein is part of the DNA-binding domain we made four variant proteins, each containing a single-amino-acid change (Fig. 1).

The DNA-binding activity of the variant proteins was measured in vitro by a DNA mobility shift assay and in vivo by complementation of a *gal4* deletion strain for growth on minimal galactose plates. By the in vitro assay, the Ala-136→Gly amino acid change reduced DNA-binding activity at least fivefold. This conclusion was supported by the complementation data. The complementation data also show that the transcriptional activation domain of the Ala-

136→Gly variant protein is functional. The reduced DNA-binding activity of this variant protein is probably due to a mutation in the DNA-binding domain. The Arg-135→Thr and the Glu-142→Lys mutations probably disrupted the DNA-binding domain, since neither variant protein bound DNA in vitro (Fig. 4) nor did they complement a *gal4*-defective strain even when the variant protein should have been overproduced. We could not directly measure overproduction, but this is likely to be the case, since the vector carrying each *LAC9* mutant was identical to the vector carrying the Ala-136→Gly mutant gene. The latter plasmid construct complemented the *gal4* defect only when present on a multicopy vector under the *ADH1* promoter, indicating that the variant LAC9 protein was overproduced. Since three of four amino acid changes reduced the binding activity of the LAC9 protein, we believe that at least a portion of the 18-27 region of the protein is part of the DNA-binding domain.

Further evidence that the 18-27 region on the carboxyl side of the zinc finger is in the DNA-binding domain comes from Johnston and Dover (11), who found that the change Thr-50→Ile in the GAL4 protein (Fig. 6) reduced DNA-binding activity.

The amino acid change Glu-140→Lys, which did not appear to affect the DNA-binding activity of LAC9, would argue that Glu-140 does not contact DNA directly. It remains to be determined whether amino acids in the 18-27 region contact DNA directly or whether they play some other role in the conformation of the DNA-binding domain. An attractive feature of this region is that it is predicted to form an α-helix (28) and could thus assume a conformation that is known to provide proteins with the ability to bind specific DNA sequences (22).

It could be argued that the 18-27 region of LAC9 and GAL4 is not actually part of the DNA-binding domain. Rather, amino acid changes in it indirectly affect the DNA-binding domain. This possibility cannot be ruled out directly at this time. However, it seems unlikely, because the Arg-135→Thr and Glu-142→Lys variant proteins are predicted to maintain the α-helical structure in the 18-27 region and should thus have a minimal effect on the conformation of the LAC9 protein. Other data argue that the conformation of the DNA-binding domain is independent of the rest of the protein sequence. For example, replacing the 807 amino acid residues on the carboxy side of residue 74 in the GAL4 protein (downstream of the 18-27 region) with the *E. coli* β-galactosidase did not destroy the DNA-binding activity of GAL4 (13). Likewise, deleting residues 75 to 767 reduced the ability of the GAL4 protein to activate transcription, but the variant protein still bound DNA (17). Finally, random mutagenesis of *GAL4* indicated that mutations affecting DNA binding cluster in the first 74 amino acids (11). Data from the bacterial regulatory proteins λ repressor, Cro, and catabolite gene activator protein also support the concept that the DNA-binding domain is independent of the other domains in a protein (22).

Similar arguments apply to the LAC9 protein, because its DNA-binding domain must have nearly the same conformation as that of GAL4. Their conformations must be similar, because they have similar amino acid sequences (Fig. 6) and the proteins activate the same genes (24, 26, 28) by binding to UAS with a common form, 5′-CGG$(N)_{11}$CCG-3′ (16).

Other mutations (11) that reduce the DNA-binding activity of GAL4 are indicated as underlined amino acids in Fig. 6. Seven of these changes fall within the zinc finger, and six changes occur in the region adjacent to the zinc finger. The latter region of the GAL4 protein is similar to those of both LAC9 and the other yeast proteins shown in Fig. 6. Such conservation of amino acids argues that the region next to the zinc finger performs a general structural function in the DNA-binding domain.

Evidence for Zn^{2+} in the GAL4 protein has been presented by Johnston (10), who showed that the Gal^- phenotype of *S. cerevisiae* carrying a mutation in the zinc finger (Pro-26→Leu) could be converted to the Gal^+ phenotype by increasing the concentration of Zn^{2+} in the culture medium. The effect was not obtained with other divalent metal ions. Because of the structural and functional homology of the LAC9 protein to the GAL4 protein, we think that the LAC9 protein has Zn^{2+} in its DNA-binding domain, but direct evidence for a Zn^{2+} requirement needs to be obtained. The Gal^- phenotype caused by the Arg-135→Thr, the Ala-136→Gly, and the Glu-142→Lys mutations in LAC9 could not be reversed by increased Zn^{2+}. However, a negative result is of little value, because most mutations in the zinc finger of the GAL4 protein, which cause a Gal^- phenotype, are not restored to a Gal^+ phenotype by Zn^{2+} (10).

Berg (1) pointed out that many proteins contain zinc fingers with the general form Cys-X_{2-4}-Cys-X_{2-15}-*a*-X_{2-4}-*a* or *a*-X_{2-4}-*a*-X_{2-15}-Cys-X_{2-4}-Cys, where *a* is Cys or His. It has been assumed that Cys or His would be functionally equivalent in zinc fingers. Our results (Fig. 5) demonstrate that this is not the case for the LAC9 protein. LAC9 proteins with a single-amino-acid change, Cys-98→His or Cys-112→His, or a double change, Cys-112→His plus Cys-115→His, did not bind DNA. A lack of DNA-binding activity could indicate that the His cannot complex to Zn^{2+} or that if it does complex to zinc the conformation of the domain is distorted. These results are in agreement with those reported by Green and Chambon (5), who showed that histidine could not functionally replace cysteine in the first zinc finger of the human estrogen receptor.

Frankel et al. (4) showed that a 30-amino-acid peptide, corresponding to the second of the nine zinc fingers in the TFIIIA protein, bound to DNA but not in the expected sequence-specific manner. Furthermore, binding did not require Zn^{2+}. The authors suggested that a single zinc finger is not able to function properly and that other residues of the protein, presumably the other zinc fingers, are needed for sequence-specific binding that requires Zn^{2+}. These data thus support our hypothesis that a single zinc finger will not confer DNA-binding specificity. Further genetic, biochemical, and structural studies will be needed to determine the role of a zinc finger and adjacent amino acids in DNA binding.

If, as we hypothesize, the DNA-binding domain of the LAC9 protein contains both a zinc finger and an α-helix motif, then this protein and ones like it may represent a transition in the evolution of DNA-binding domains from procaryotes with a helix-turn-helix motif to higher eucaryotes with multiple zinc fingers.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant GM-22749 from the National Institutes of Health and by Biomedical Support Grant RR05374 from the Biomedical Research Support Branch, Division of Research Facilities and Resources, National Institutes of Health, to R.C.D. Computer resources used to carry out these studies were provided by the BIONET National Computer Resource for Molecular Biology, whose funding is provided by the Biomedical Research Technology Program, Division of Research Resources, National Institutes of Health, grant RR 01865-05.

We thank Gary Hellmann for expert assistance with in vitro translation experiments and M. S. Coleman, J. Davidson, and J. Lesnaw for valuable discussions and critical reading of the manuscript.

LITERATURE CITED

1. **Berg, J. M.** 1986. Potential metal-binding domains in nucleic acid binding proteins. Science **232**:485–487.
2. **Blumberg, H., A. Eisen, A. Sledziewski, D. Bader, and E. T. Young.** 1987. Two zinc fingers of a yeast regulatory protein shown by genetic evidence to be essential for its function. Nature (London) **328**:443–445.
3. **Chen, E. Y., and P. H. Seeburg.** 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA **4**: 165–170.
4. **Frankel, A. D., J. M. Berg, and C. O. Pabo.** 1987. Metal-dependent folding of a single zinc finger from transcription factor IIIA. Proc. Natl. Acad. Sci. USA **84**:4841–4845.
5. **Green, S., and P. Chambon.** 1987. Oestradiol induction of a glucocorticoid-responsive gene by a chimeric receptor. Nature (London) **325**:75–78.
6. **Hanas, J. S., D. J. Hazuda, D. F. Bogenhagen, F. Y.-H. Wu, and C.-W. Wu.** 1983. *Xenopus* transcription factor A requires zinc for binding to the 5S RNA gene. J. Biol. Chem. **258**:1420–1425.
7. **Hartshorne, T. A., H. Blumberg, and E. T. Young.** 1986. Sequence homology of the yeast regulatory protein ADRI with *Xenopus* transcription factor TFIIIA. Nature (London) **320**: 283–287.
8. **Hellmann, G. M., S. T. Hiremath, J. G. Shaw, and R. E. Rhoads.** 1986. Cistron mapping of tobacco vein mottling virus. Virology **151**:159–171.
9. **Hope, I. A., and K. Struhl.** 1985. GCN4 protein, synthesized in vitro, binds *HIS3* regulatory sequences: implications for general control of amino acid biosynthetic genes in yeast. Cell **43**:177–188.
10. **Johnston, M.** 1987. Genetic evidence that zinc is an essential co-factor in the DNA binding domain of the GAL4 protein. Nature (London) **328**:353–355.
11. **Johnston, M., and J. Dover.** 1987. Mutations that inactivate a yeast transcriptional regulatory protein cluster in an evolutionarily conserved DNA binding domain. Proc. Natl. Acad. Sci. USA **84**:2401–2405.
12. **Kammerer, B., A. Guyuonvarch, and J. C. Hubert.** 1984. Yeast regulatory gene PPR1. I. Nucleotide sequence, restriction map, and codon usage. J. Mol. Biol. **180**:239–250.
13. **Keegan, L., G. Gill, and M. Ptashne.** 1986. Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. Science **231**:699–704.

13a. **Kunkel, T. A., J. D. Roberts, and R. A. Zakour.** 1987. Rapid and efficient site-specific mytagenesis without phenotypic selection. Methods Enzymol. **154**:367–382.

14. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
15. **Laughton, A., and R. F. Gesteland.** 1984. Primary structure of the *Saccharomyces cerevisiae* GAL4 gene. Mol. Cell. Biol. **4**: 260–267.
16. **Leonardo, J. M., S. M. Bhairi, and R. C. Dickson.** 1987. Identification of upstream activator sequences that regulate induction of the β-galactosidase gene in *Kluyveromyces lactis*. Mol. Cell. Biol. **7**:4369–4376.
17. **Ma, J., and M. Ptashne.** 1987. Deletion analysis of *GAL4* defines two transcriptional activating segments. Cell **48**:847–853.
18. **Marsh, J. L., M. Erfle, and E. J. Wykes.** 1984. The pIC plasmid and phage vectors with versatile cloning sites for recombinant selection by insertional inactivation. Gene **32**:481–485.
19. **Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12**:7035–7056.
20. **Messenguy, F., E. Dubois, and F. Descamps.** 1986. Nucleotide sequence of the ARGRII regulatory gene and amino acid sequence homologies between ARGRII, PPRI, and GAL4 regulatory proteins. Eur. J. Biochem. **157**:77–81.
21. **Miller, J., A. D. McLachlan, and A. Klug.** 1985. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. EMBO J. **4**:1609–1614.
22. **Pabo, C. O., and R. T. Sauer.** 1984. Protein-DNA recognition. Annu. Rev. Biochem. **53**:293–321.
23. **Riley, M. I., and R. C. Dickson.** 1984. Genetic and biochemical characterization of the galactose gene cluster in *Kluyveromyces lactis*. J. Bacteriol. **158**:705–712.
24. **Riley, M. I., J. E. Hopper, S. A. Johnston, and R. C. Dickson.** 1987. *GAL4* of *Saccharomyces cerevisiae* activates the lactose-galactose regulon of *Kluyveromyces lactis* and creates a new phenotype: glucose repression of the regulon. Mol. Cell. Biol. **7**: 780–786.
25. **Ruzzi, M., K. D. Breunig, A. G. Ficca, and C. P. Norrenberg.** 1987. Positive regulation of the β-galactosidase gene from *Kluyveromyces lactis* is mediated by an upstream activation site that shows homology to the GAL upstream activation site of *Saccharomyces cerevisiae*. Mol. Cell. Biol. **7**:991–997.
26. **Salmeron, J. M., Jr., and S. A. Johnston.** 1986. Analysis of the *Kluyveromyces lactis* positive regulatory gene *LAC9* reveals functional homology to, but sequence divergence from, the *Saccharomyces cerevisiae GAL4* gene. Nucleic Acids Res. **14**: 7767–7781.
27. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
28. **Wray, L. V., M. M. Witte, R. C. Dickson, and M. I. Riley.** 1987. Characterization of a positive regulatory gene, *LAC9*, that controls induction of the lactose-galactose regulon of *Kluyveromyces lactis*: structural and functional relationships to GAL4 of *Saccharomyces cerevisiae*. Mol. Cell. Biol. **7**:1111–1121.
29. **Zhou, K., P. R. G. Brisco, A. E. Hinkkanen, and G. B. Kohlhaw.** 1987. Structure of yeast regulatory gene *LEU3* and evidence that *LEU3* itself is under general amino acid control. Nucleic Acids Res. **15**:5261–5273.