MOLECULAR AND CELLULAR BIOLOGY, Apr. 1991, p. 1777–1784
0270-7306/91/041777-08$02.00/0
Copyright © 1991, American Society for Microbiology
Vol. 11, No. 4

Identification of Base and Backbone Contacts Used for DNA Sequence Recognition and High-Affinity Binding by LAC9, a Transcription Activator Containing a C6 Zinc Finger

YUAN-DI C. HALVORSEN, KRISHNAN NANDABALAN, AND ROBERT C. DICKSON*

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

Received 5 October 1990/Accepted 3 January 1991

The LAC9 protein of *Kluyveromyces lactis* is a transcriptional regulator of genes in the lactose-galactose regulon. To regulate transcription, LAC9 must bind to 17-bp upstream activator sequences (UASs) located in front of each target gene. LAC9 is homologous to the GAL4 protein of *Saccharomyces cerevisiae*, and the two proteins must bind DNA in a very similar manner. In this paper we show that high-affinity, sequence-specific binding by LAC9 dimers is mediated primarily by 3 bp at each end of the UAS: $\begin{matrix}-8\text{CGG}(\text{N}_{11})\text{CCG}+8 \\ +8\text{GCC}(\text{N}_{11})\text{GCC}-8\end{matrix}$. In addition, at least one half of the UAS must have a GC or CG base pair at position 1 for high-affinity binding; LAC9 binds preferentially to the half containing the GC base pair. Bases at positions 2, 3, and 4 in each half of the UAS make little if any contribution to binding. The center base pair is not essential for high-affinity LAC9 binding when DNA-binding activity measured in vitro. However, the center base pair must play an essential role in vivo, since all natural UASs have 17, not 16, bp. Hydroxyl radical footprinting shows that a LAC9 dimer binds an unusually broad region on one face of the DNA helix. Because of the data, we suggest that LAC9 contacts positions 6, 7, and 8, both plus and minus, of the UAS, which are separated by more than one turn of the DNA helix, and twists part way around the DNA, thus protecting the broad region of the minor groove between the major-groove contacts.

Transcription of genes is enhanced or repressed by regulatory proteins that recognize and bind to specific DNA sequences or *cis*-acting regulatory elements. The molecular basis of such protein-DNA recognition processes is only beginning to be understood, and much remains to be elucidated, particularly since it is clear that eucaryotes have evolved many unique protein domains for recognizing specific DNA sequences. One type of domain, found in 11 different fungal transcription activators, contains six conserved cysteines with the form C-aa_2-C-aa_6-C-aa_{6-9}-C-aa_2-C-aa_6-C, where aa is any amino acid (reviewed in reference 1). Two members of this family of proteins, GAL4 (16) and LAC9 (11), have been shown to bind two zinc atoms, presumably by coordination of the six cysteines (16). For these and other historical reasons, this family of proteins has been referred to as having a C6 zinc finger (9). LAC9 and GAL4 are homologous, and they must bind DNA is a similar manner (17, 19). To further our understanding of how the proteins in this family recognize a specific DNA sequence, we have determined the bases that specify binding by LAC9.

LAC9 enhances transcription of genes in the lactose-galactose regulon in the yeast *Kluyveromyces lactis* (28). Transcriptional enhancement of the regulon is controlled so that in the absence of inducer there is low expression, while in the presence of an inducer, such as lactose (7), expression is high. In order to enhance transcription of the target genes, LAC9 must bind to a 17-bp element, termed an upstream activating sequence (UAS), having the consensus form 5′-CGG(N_5)(A/T)(N_5)CCG-3′ (3, 13, 18). The UASs are located upstream of each gene in the regulon and are always grouped in two or four copies (summarized in reference 8).

The DNA-binding domain of LAC9 is located between amino acid residues 85 and 228 (11), a region that contains the C6 zinc finger as well as amino acids on the carboxyl side of the finger region. LAC9 variants with single amino acid changes in the zinc finger or a nearby region decrease DNA-binding activity, suggesting that the DNA-binding domain is composed of at least two regions (26). Recently we showed that DNA-binding specificity is determined in part by a region of 14 amino acids adjacent to the carboxyl side of the C6 zinc finger of LAC9 (27). Corton and Johnston (6) obtained a similar result with GAL4. In addition, we showed that the C6 zinc finger plays an essential role in the binding specificity and affinity of LAC9 (27). However, the exact role(s) of the C6 zinc finger in DNA binding remains to be determined.

To begin to understand how LAC9 contacts a UAS, we have quantitated the contribution each base in the UAS makes to DNA-binding activity. In addition, base contacts between the UAS and LAC9 were determined by missing-contact footprinting (4), and deoxyribose or backbone contacts were determined by hydroxyl radical footprinting (25). The data show that the highly conserved, terminal 3 bp at both ends of the UAS are essential for binding by LAC9, as are a few of the other bases; most of the other bases contribute little to binding specificity or affinity.

MATERIALS AND METHODS

LAC9 protein. The LAC9 DNA-binding domain, termed LAC9(85-228*), was produced in *Escherichia coli* and purified to greater than 98% homogeneity (11). Because of the recombinant plasmid used for expression in *E. coli*, LAC9(85-228*) contained a methionine at its N terminus and four extra amino acids (PSDL) at its C terminus. These extra amino acids did not interfere with DNA-binding activity (11). A molar extinction coefficient at 280 nm of 2.24×10^4 cm^{-1}

* Corresponding author.

TABLE 1. UAS sequences

Genetic locus	Name of sequence	Sequence
GAL1	UASI	CGGATCACTTCAATCCG
	UASII	CGGATTGAACGAATCCG
GAL10	UASI	CGGAATACACCAATCCG
	UASII	CGGAAGACTATGCCCCG
GAL7	UASI	CGGAGTTGAGTGGTCCG
	UASII	CGGATGATTGTGGTCCG
LAC4	UASI	CGGAAATTTGTGGTCCG
	UASII	CGGAATTCTGTTCACCG
LAC12	UASI	CGGGATTCTTCCCTCCG
	UASII	CGGCCCACTGACTCCCG
Consensus[a]		CGGAATACTGTNNTCCG

[a] To be considered in the consensus, a base had to occur in more than 50% of the individual UASs. References for the UAS sequences are in Dickson and Riley (8).

was determined for LAC9(85-228*) from amino acid analysis (11), and this value was used to determine protein concentrations. The fraction of protein capable of binding DNA (active protein) was determined by titrating a constant concentration of LAC9(85-228*) with increasing concentrations of the consensus (cUAS) DNA: 30 to 40% of LAC9(85-228*) was active and bound DNA (11).

cUAS and mutant UAS DNA. The 17-bp cUAS, 5′-CGGAATACTGTATTCCG-3′, was deduced from the natural UASs of the lactose-galactose regulon of *K. lactis* (Table 1). The cUAS was designed to have dyad symmetry around but exclusive of the center base so that each of the other bases could be systematically and symmetrically changed. Bases in the upper strand of the cUAS (the strand with a T in the center) to the left of center are denoted by negative numbers, and those to the right of center are denoted by positive numbers (Table 2). The cUAS and mutant UASs were synthesized on an Applied Biosystems 380B automated DNA synthesizer. In all cases, an oligonucleotide was self-complementary except for the central base, which was either A and T. Thus each oligonucleotide self-annealed to form a double-stranded, blunt-ended DNA. UASs were cloned into the *Eco*RV site of pIC20R (14). The presence of a single UAS in a plasmid was checked by dideoxynucleotide sequencing. For DNA-binding assays, the UAS was removed from the vector as a 101-bp *Eco*RI fragment. The same *Eco*RI vector fragment lacking a UAS was used to determine nonspecific DNA binding. UASI of the *LAC4* promoter, termed LAC4UASI (13), was synthesized as a 22-bp oligonucleotide and cloned in the *Sal*I site of pIC20R. A 108-bp *Eco*RI fragment was isolated from this plasmid and used to determine the relative K_d of LAC4UASI.

Radioactive labeling of DNA. To determine the relative dissociation constants (K_ds) of the cUASs and mutant UASs, the 101-bp *Eco*RI fragment from pIC20UAS was

TABLE 2. UAS binding affinity

Sequence designation	UAS sequence																	Relative K_d[a]	ΔΔG (kcal/mol)
	−8	−7	−6	−5	−4	−3	−2	−1	0	+1	+2	+3	+4	+5	+6	+7	+8		
cUAS	C	G	G	A	A	T	A	C	T	G	T	A	T	T	C	C	G	1 ± 0.1	0.00
−8G/+8C	G																C	207 ± 7.1	−3.16
−8A/+8T	A																T	7 ± 0.2	−1.17
−8T/+8A	T																A	184 ± 7.3	−3.13
−7C/+7G		C														G		330[b] ± 7.5	−3.48
−7A/+7T		A														T		330[b] ± 7.5	−3.48
−7T/+7A		T														A		330[b] ± 7.5	−3.48
−6C/+6G			C												G			12 ± 0.3	−1.48
−6A/+6T			A												T			185 ± 2.0	−3.13
−6T/+6A			T												A			129 ± 4.7	−2.91
−5T/+5A				T										A				0.4 ± 0.03	+0.55
−5C/+5G				C										G				25 ± 0.6	−1.93
−5G/+5C				G										C				17 ± 1.9	−1.71
−4T/+4A					T								A					1 ± 0.3	−0.06
−4C/+4G					C								G					3 ± 0.2	−0.64
−4G/+4C					G								C					1 ± 0.08	−0.20
−3A/+3T						A						T						2 ± 0.07	−0.47
−3C/+3G						C						G						2 ± 0.3	−0.45
−3G/+3C						G						C						135 ± 2.6	−2.94
−2T/+2A							T				A							0.5 ± 0.06	+0.42
−2C/+2G							C				G							6 ± 0.4	−1.03
−2G/+2C							G				C							0.8 ± 0.06	+0.13
−1G/+1C								G		C								1 ± 0.2	−0.16
−1A/+1T								A		T								162 ± 2.6	−3.05
−1T/+1A								T		A								314 ± 8.0	−3.45
0/A									A									1 ± 0.4	−0.11
0/G									G									4 ± 0.3	−0.83
0	C	G	G	A	A	T	A	C	–	G	T	A	T	T	C	C	G	8 ± 0.3	−1.25
LAC4UASI	C	G	G	A	A	A	T	T	T	G	T	G	G	T	C	C	G	4 ± 0.4	−0.83
LAC4UASII	C	G	G	A	A	T	T	C	T	G	T	T	C	A	C	C	G	3 ± 0.3	−0.60
LAC9UAS	C	G	G	A	C	C	G	G	A	A	C	A	G	A	C	C	G	5 ± 0.8	−1.01

[a] Each value represents the average of three or more determinations ± 1 standard deviation. The K_d for the cUAS was $(3.1 \pm 0.4) \times 10^{-10}$ M when determined by using a DNA concentration of 6×10^{-10} M, but nearly identical values were obtained with DNA concentrations of 1.5×10^{-10} and 3×10^{-10} M.

[b] This value reflects the relative K_d of LAC9(85-228*) for nonspecific DNA and thus the sensitivity of the assay.

radiolabeled by filling in the 5′ unpaired ends using [α-^{32}P] dATP and the Klenow fragment of DNA polymerase I (Pharmacia LKB, Piscataway, N.J.). A specific activity of 10^7 cpm/pmol was usually obtained. A 108-bp *Eco*RI fragment containing LAC4UASI was used to determine relative dissociation constants (K_ds). UASII was obtained as a 247-bp *Sna*BI-to-*Nde*I fragment from the promoter region of *LAC4* (from nucleotide +354 to −916; 13) and 5′ end labeled using [γ-^{32}P]ATP and T4 polynucleotide kinase. LAC9UAS was obtained as a 417-bp *Acc*I-to-*Hin*dIII fragment from the promoter region of *LAC9* (nucleotides 376 to 792; 28) and 5′ end labeled at the *Hin*dIII site as described above for cUAS labeling.

DNA used for footprint analysis was prepared by digesting pICcUAS with *Asp* 718 and either 5′ end labeled by using T4 polynucleotide kinase and [γ-^{32}P]ATP or 3′ end labeled by using the Klenow fragment of DNA polymerase I and [α-^{32}P] dGTP. The resulting radioactive DNA was cleaved with *Eco*RI and electrophoresed on a 2% agarose gel, and the 89-bp fragment carrying the UAS was electroeluted onto a DEAE membrane (2). A specific activity of 1.5×10^6 cpm/pmol was usually obtained.

Nitrocellulose filter-binding assay. The affinity of LAC9(85-228*) for each UAS was assayed by a nitrocellulose filter-binding assay. A ^{32}P-labeled UAS-containing DNA fragment (6×10^{-10} M) was incubated with various amounts of LAC9 (85-228*) in a total of 45 μl of binding buffer (20 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid [HEPES; pH 7.9], 100 mM NaCl, 7 mM β-mercaptoethanol, 0.1 mM EDTA, 5% glycerol) at room temperature for 5 min. Samples were then filtered through a nitrocellulose filter (Schleicher & Schuell; BA-85, 0.22-μm pore size) in approximately 5 s by using a slight vacuum. Background radioactivity retained on the filters was determined in an identical binding reaction without LAC9(85-228*). Filters were dried and counted in a liquid scintillation counter. Total radioactivity was determined by spotting DNA on a filter. After the background radioactivity was subtracted, the amount of the protein-DNA complex retained on the filter was calculated as a percentage of total DNA. The dissociation constant (K_d) of cUAS was determined by using a constant concentration of DNA (6×10^{-10} M) and increasing concentrations of LAC9(85-228*). The K_d was calculated by using the formula $K_d = [C][D]/[CD]$, where $[C]$ is the concentration of the free protein, $[D]$ is the concentration of the free DNA, and $[CD]$ is the concentration of the protein-DNA complex when 50% of the DNA was bound. The concentration of LAC9(85-228*) represents that of a dimer, since LAC9 binds DNA as a dimer (11). Relative K_ds were calculated by using the formula $K_m/K_c = (C_m/D_m)/(C_c/D_c)$, where C is the concentration of the protein-DNA complex and D is the concentration of the free cUAS (c) or mutant UAS (m). Relative K_ds were determined from at least three different protein concentrations, each assayed in duplicate. The free energy change, $\Delta\Delta G$, was calculated by using the formula $\Delta\Delta G = -RT \ln(\text{relative } K_d)$, where R is 1.987 cal mol^{-1} K^{-1} (1 cal = 4.184 J) and T is temperature (298 K) (24).

Footprint analysis. Missing-contact footprints were obtained by using the procedure of Brunelle and Schleif (4). Generally, 10^5 cpm of end-labeled and chemically premodified DNA was incubated with enough LAC9(85-228*) to bind approximately 30% of the DNA. The protein-bound and free DNAs were separated by electrophoresis (10) on a 6% polyacrylmide gel (acrylamide-bisacrylamide, 29:1) and identified by autoradiography. The radioactive bands representing free and bound DNA were excised from the gel and electroeluted onto a DEAE membrane (2). The DNA backbone was cleaved at the premodified bases by treatment with piperidine. The quantity of fragments in the bound- and free-DNA samples were analyzed on a 10% polyacrylamide–8 M urea DNA sequencing gel.

Methylation interference footprints were prepared by using the procedure of Sienbenlist and Gilbert (23).

Hydroxyl radical footprints were obtained by using a modification of the procedure of Tullius and Dombroski (25). Briefly, ^{32}P-end-labeled DNA (10^5 cpm) was incubated with LAC9(85-228*) for 5 min at room temperature in standard binding buffer lacking glycerol. The DNA was treated with hydroxyl radicals, and the reaction was stopped after 2 min by adding thiourea to 166 mM. The sample was immediately subjected to electrophoresis on a 6% native polyacrylamide gel to separate free and bound DNAs. DNA samples were purified as described above for missing-contact footprinting and analyzed on a 10% polyacrylamide–8 M urea gel. DNA size markers were produced by chemical fragmentation of ^{32}P-labeled DNA (15).

RESULTS

Affinity of LAC9 for UAS. The ability of LAC9 to activate transcription is dependent on binding of the protein to UAS. Evidence supporting this idea comes mainly from studies of the two UASs in front of *LAC4* (3, 13, 18). These two UASs have slightly different sequences (Table 1), which indicates that some, but not all, of the 17 bp in the UAS are essential for LAC9 binding. To begin to identify the essential bases, we compared the base sequences of the known and predicted LAC9 UASs (Table 1). Three important generalizations can be made from the sequence comparison. First, the two halves of the UAS are almost identical and are related by twofold rotational symmetry about the center base. This suggests that LAC9 binds the UAS as a dimer or an even-numbered multimer. Physical evidence for a dimer has been obtained (11). Second, the three bases at the end of each half of the UAS are conserved, and hence it is reasonable to assume that these bases must be essential for LAC9 binding. Third, a consensus sequence having perfect twofold rotational symmetry about the center base can be derived and the function of each base in the UAS can be analyzed by starting with a symmetrical cUAS (Table 1). We verified this assumption by showing that the K_d of the cUAS, 3×10^{-10} M, was nearly the same as that of the natural UAS (Table 2); in fact, the cUAS had the highest affinity for LAC9.

To examine the contribution of each base pair in the UAS to LAC9 binding, we examined the effect of pairwise, symmetrical base changes on DNA-binding affinity. Changes in binding were quantified by using a filter-binding assay. Binding affinities for all mutant UASs were determined and expressed as relative dissociation constants compared with the K_d of cUAS (Table 2).

As predicted, changes in any of the three conserved bases at the ends of the UAS greatly reduced affinity for LAC9. Of these bases, the −7 and +7 positions were the most sensitive to change, since substitution with any of the other three bases resulted in a complete loss of sequence-specific binding by LAC9. A large reduction in binding activity was produced by two of the three base changes, i.e., those at −8 and +8 and at −6 and +6.

Changes in the center base of the UAS, which is either an A or a T in all natural UASs, did not significantly affect LAC9 binding, and deletion of the center base reduced

binding only eightfold. Both results indicate that the center base does not make a sequence-specifying contact.

Missing-contact and methylation interference footprint analyses. Missing-contact footprinting was used to determine which bases in the UAS contact LAC9. This technique is very sensitive and, in contrast to the filter-binding assay, is able to reveal differences in contacts between the two halves of the UAS. In this procedure, the UAS-containing DNA is lightly depurinated or depyrimidinated so that on average only one base is removed in each strand of DNA (4). The DNA is then incubated with LAC9, and the sample is electrophoresed on a gel to separate the protein-DNA complex, referred to as the bound sample, from the uncomplexed, free DNA. The backbones of the DNAs in the bound- and free-DNA samples are cleaved wherever a base has been removed. The sizes and quantities of the DNA products are revealed on a DNA sequencing gel following autoradiography. If the removal of a particular base from the UAS prevents binding of the protein, then the free-DNA sample will show enhancement of a DNA fragment and the corresponding fragment in the bound DNA sample will be reduced, indicating contact of the protein with that particular base. A similar rationale applies to methylation interference footprinting, except that reduced binding is due to methylation of Gs or As: G is methylated at N-7 in the major groove and A is methylated at N-3 in the minor groove of the DNA.

Missing-contact footprints of the cUAS are shown in Fig. 1. The results are displayed for each strand of DNA; the strands have the same base sequence except for the center base, which is a T or an A as indicated in the figure. Footprints were analyzed by comparing band intensities in the bound- and free-DNA samples. For example, in the depurination reaction (G+A) for the strand with a T in the center, there is a large decrease in LAC9 binding (the radioactive band shifts from the bound- to the free-DNA sample) when the −6G is missing, a medium decrease when −7G is missing, and a lesser decrease when +1G or +8G is missing. For the depyrimidination reaction (C+T), a large decrease in LAC9 binding is seen when +2T is missing, and a medium decrease is seen when −8C, the center T, +4T, +5T, +6C, or +7C is missing. The loss of LAC9 binding due to removal of Cs agrees with the filter-binding assay data.

Also shown in Fig. 1 are the results of a methylation interference footprint (labeled G) in which methylation of −7G and −6G decreases the amount of protein-bound DNA, a result predicted from the filter-binding data.

The results of these experiments are summarized and diagrammed at the bottom of Fig. 1. Their significance and relationship to filter-binding and hydroxyl radical footprint data are presented in the Discussion.

Missing-contact and methylation interference footprinting of mutant UAS. All natural UASs (Table 1) have CGG at their 5′ ends and CCG at their 3′ ends, which suggests that these bases are necessary for DNA-binding activity. This idea was supported by the data obtained using the filter-binding assay and mutant UAS containing symmetrical base changes, with the exception of two mutant UAS, −8A/+8T and −6C/+6G, that unexpectedly bound LAC9 fairly well (Table 2). To try and understand how these mutant UASs bound LAC9, we performed a missing-contact footprint analysis. Experimental data and a diagrammatic summary are shown in Fig. 2. Two generalizations can be made from the data. First, the base contacts made by LAC9 with cUAS are also made with the two mutant UASs except for the mutant base pairs. Second, LAC9 makes additional base contacts that are unique for each mutant UAS.

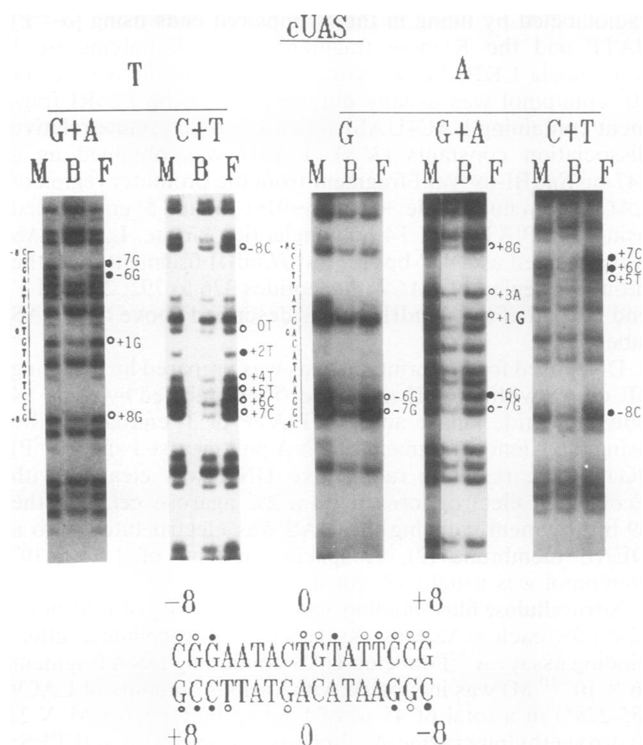

FIG. 1. Missing-contact and methylation interference footprint analyses of cUAS. Footprints were obtained by using LAC9(85-228*) and a DNA fragment carrying the cUAS. The DNA fragment was end labeled in the strand having either a T as the center base (gels marked T) or an A (gels marked A). DNAs were modified by depurination (G+A), depyrimidation (C+T), or methylation (G) before LAC9(85-228*) was added. B, DNA bound to LAC9; F, free or unbound DNA; M, molecular length markers. The sequence and orientation of the UAS are indicated to the left of the footprints. Strong base contacts to LAC9 are indicated by a full circle, and weak contacts are indicated by an open circle. The data are summarized in the bottom diagram.

Hydroxyl radical footprint analysis. Hydroxyl radical footprinting is a high-resolution technique for determining the deoxyribose backbone positions in close contact with a bound protein (25). As shown in Fig. 3A, LAC9(85-228*) protected the DNA against cleavage at positions close to the center of the UAS. Weak but significant protection was also found at some positions near the ends of the UAS. The protection pattern was analyzed by scanning the bound and free DNA (Fig. 3B), and the results are summarized in Fig. 3C.

DISCUSSION

To begin to understand how LAC9, a member of a family of fungal proteins characterized by the presence of a C6 zinc finger domain, binds to DNA, we analyzed the base and backbone contacts the protein makes with its cognate UAS. The data are summarized in Fig. 4. The general conclusions are that (i) the three highly conserved base pairs at each end of the 17-bp UAS supply crucial contacts for sequence recognition and high-affinity binding by LAC9; (ii) the two halves of the UAS are contacted in a similar but not identical manner by LAC9; (iii) a LAC9 dimer binds one face of the DNA so that each monomer contacts an adjacent major groove, with the intervening minor groove in close contact

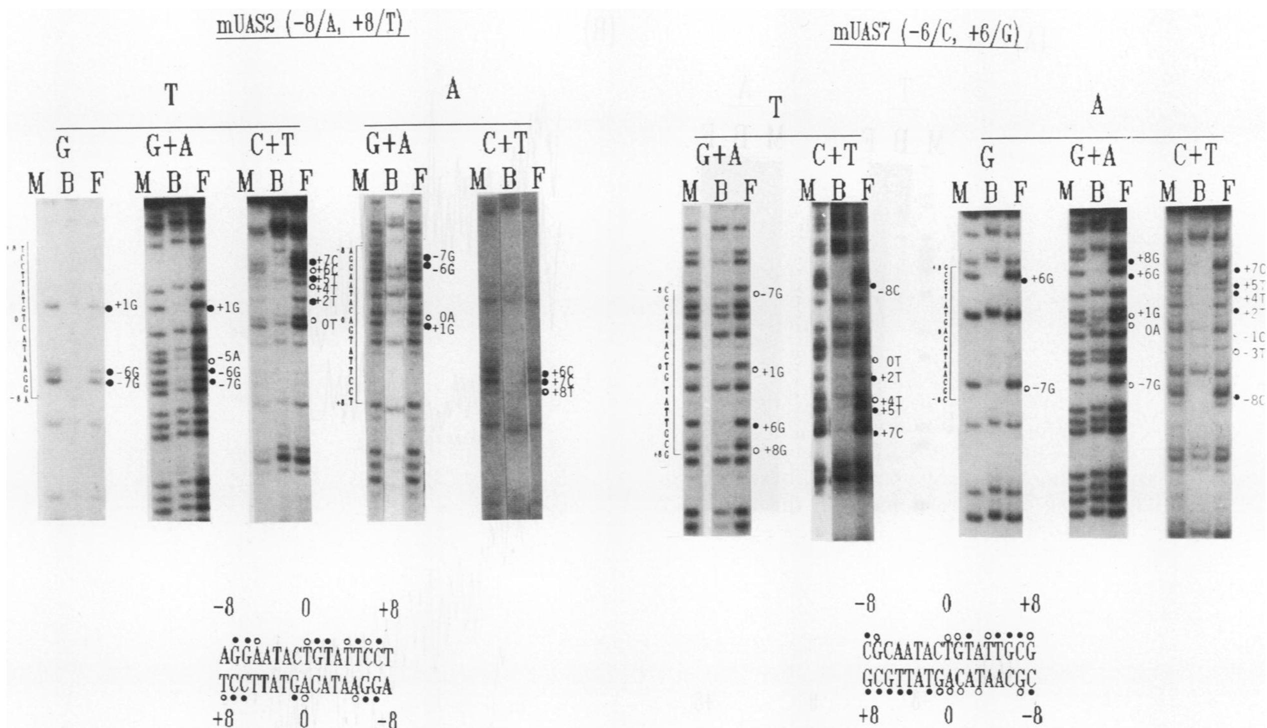

FIG. 2. Missing-contact and methylation interference footprints of mutant UAS (mUAS). Base contacts made by LAC9(85-228*) with the mutant UAS having the symmetrical changes −8A/+8T and −6C/+6G were determined. Symbols and abbreviations are defined in the legend to Fig. 1. Data are summarized in the bottom diagram.

with the protein; and (iv) LAC9 and GAL4 bind DNA in a very similar manner, which suggests that related fungal proteins with a C6 zinc finger motif use a similar strategy to bind DNA.

Because all known LAC9 UASs (Table 1) have the sequence 5′-CGG(N_5)(A/T)(N_5)CCG-3′, we expected any base change in the three terminal bases to significantly reduce affinity for LAC9. Only the +7/−7 base pair followed this prediction, since changing this base pair to any other base pair caused complete loss of LAC9 binding as measured by the filter-binding assay (Table 2). In addition, missing contact footprinting showed that loss of either base at the −7 or +7 position reduced LAC9 binding (Fig. 1). Finally, LAC9 binding was reduced by methylation of the −7G on the strand of the cUAS having an A in the center (Fig. 1). We conclude that the −7/+7 base pairs form sequence-specifying, high-affinity contacts with LAC9. Furthermore, particularly because of the methylation interference data, we predict that −7G in each half of the UAS contacts a specific LAC9 amino acid residue(s). The missing contact data argue that +7C in each half of the UAS also contacts LAC9. As used here and throughout the text, the term contact represents an inference made from the data and does not necessarily represent a direct, noncovalent interaction. Other explanations for loss of DNA binding, such as conformational distortion in the DNA due to removal of a base (4), are possible.

In contrast to the data for the −7/+7 position, only two of three base pair changes at position −8/+8 significantly reduced LAC9-binding activity (Table 2). Changing from −8C/+8G to −8A/+8T reduced binding only sevenfold. The ability of LAC9 to tolerate a T instead of a G at +8 may be due to the presence on either base of a hydrogen bond acceptor protruding into the major groove (21). This possibility is supported by the methylation interference footprint (Fig. 1), since methylation of N-7 on the +8G residue did not reduce LAC9 binding. Such toleration of a methyl group protruding into the major groove would allow substitution with a T whose C-5 methyl group would have roughly the same spatial coordinates as the N-7 methyl of G. Another explanation for the ability of A to substitute for C at position −8 may be the presence on either base of a hydrogen bond donor protruding into the major groove (21). In addition to the filter-binding data, missing-contact footprint data show that LAC9 binding depends on the presence of each base at the ±8 positions in the UAS (Fig. 2). The way in which LAC9 is able to compensate for the loss in free energy produced by the change from −8C/+8G to −8A/+8T is revealed by the missing-contact footprint data shown in Fig. 2: free-energy compensation is achieved by making contact with +8T in one half of the UAS, with −4A in both halves of the UAS, with +2T in one half of the UAS, and with +4T and +5T in one half of the UAS. We conclude that LAC9 makes sequence recognition, high-affinity contacts with the −8 and +8 positions in the UAS.

The missing-contact footprints of the cUAS (Fig. 1) indicate the importance of −6G and +6C for normal, high-affinity LAC9 binding, as does the methylation interference data (Fig. 1), which show that methylation of −6G on the A-containing strand of the cUAS interferes with high-affinity LAC9 binding. On the basis of this result and the large loss in free energy of binding (Table 2) associated with a base change in the −6G position, we conclude that it is quite likely that an amino acid(s) in LAC9 makes a hydrogen bond contact(s) to this G in both halves of the UAS. Combining the data presented in Table 2 and Fig. 1 and 2, we conclude

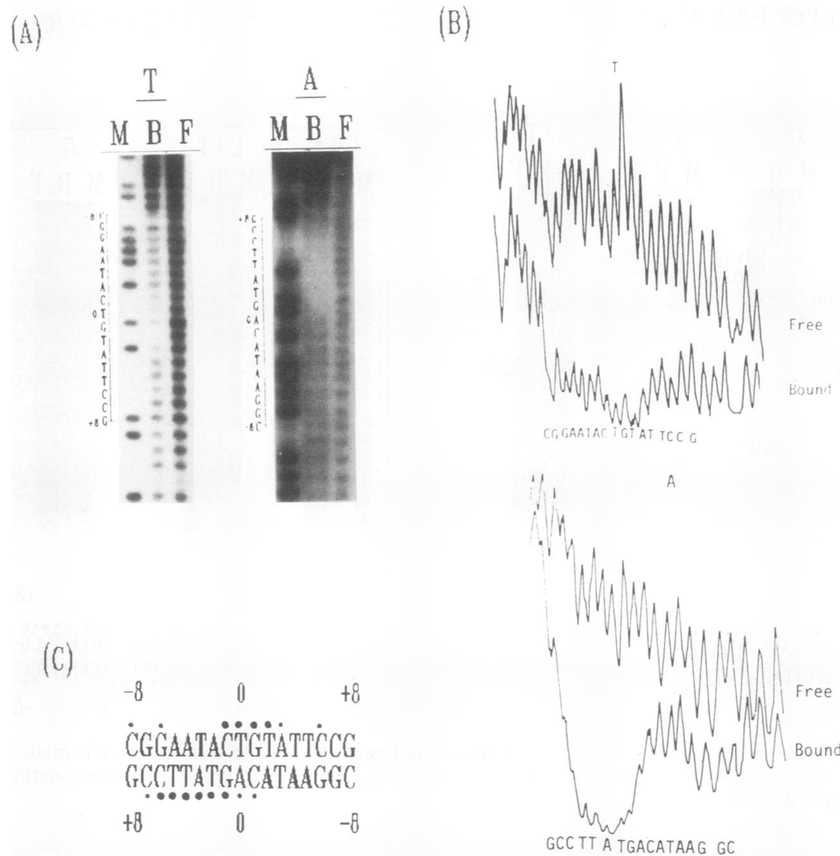

FIG. 3. Hydroxyl radical footprint analysis of LAC9(85-228*) bound to cUAS. (A) Hydroxyl radical footprints are shown for each strand of the cUAS, one strand having a T and the other an A at the center base. M, B, and F represent molecular size standards, DNA bound by LAC9, and free DNA, respectively. Base numbering for the UAS is shown at the side of the autoradiograms. (B) Densitometer scans of the footprints shown in panel A. (C) Diagrammatic summary of the data showing deoxyriboses that are strongly (●) and moderately (○) protected from hydroxyl radical attack.

that the −6 and +6 bases are necessary for sequence-specific, high-affinity LAC9 binding.

The symmetrical base changes −6G/+6C to −6C/+6G reduced LAC9 binding 12-fold as measured by the filter-binding assay. A simple explanation for this result based on hydrogen bond donors or acceptors is not obvious. The missing-contact footprinting data for the mutant UAS carrying this symmetrical change offers a partial explanation for LAC9 binding to mutant UAS7. The footprints of mutant UAS7 (Fig. 2) show that LAC9 compensates for the −6/+6 base pair change by contacting the two +6C residues (they are Gs in the cUAS) and by increasing contact with other bases. Thus, at least for this UAS, LAC9 is able to partially compensate for the loss in free energy caused by the mutant bases.

At least one other base pair, at position 1, is necessary for high-affinity LAC9 binding (Table 2 and Fig. 1). LAC9 does not contact the N-7 position of +1G in the A-containing strand of the cUAS because methylation of this site does not interfere with binding. However, as discussed below, LAC9 binds more strongly to one half of the UAS than to the other, so binding to the cUAS may be normal as long as one of the +1Gs is not methylated. Why only a GC or CG base pair is allowed at position 1 is not obvious. This lack of discrimination suggests that the base pair at position 1 does not contribute to sequence recognition. Two natural UASs, LAC4UASI and LAC9UAS, have a GC or CG base pair in

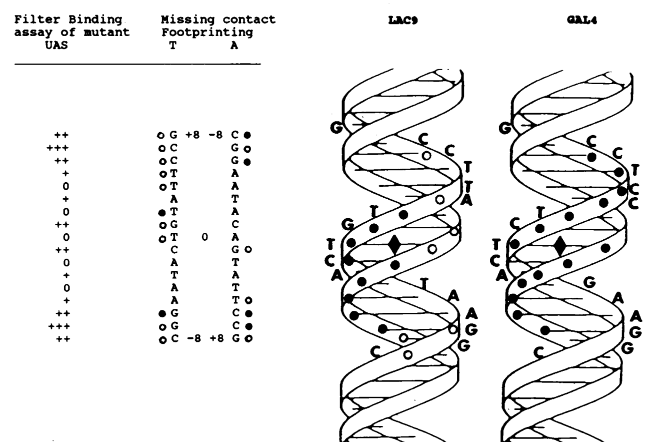

FIG. 4. Diagrammatic summary of data from the filter-binding assay. +++, All base changes decreased binding; ++, some base changes decreased binding; +, one base change decreased binding; 0, binding not affected by any base change. Data from missing-contact footprinting are indicated: ●, strong base contact; ○, moderate base contact. The results of hydroxyl radical footprinting are presented on a ribbon model of B-form DNA. A diamond marks the center base. Circles on the ribbon indicate bases whose deoxyriboses were strongly (●) or moderately (○) protected by bound LAC9 from attack by hydroxyl radical. The data for GAL4 are from Carey et al. (5).

one half of the UAS only, indicating that binding is probably asymmetric, with the half containing the GC pair binding most strongly.

The ±5 positions make a small contribution to LAC9-binding affinity according to both filter-binding and missing-contact footprint data. Since either an A or a T is tolerated at this position, there is likely to be little if any sequence specificity contributed by this base pair.

The ±2 position does not contribute to LAC9 binding, since any base pair gives reasonably strong binding activity. Removal of +2T by the missing-contact technique reduced LAC9 binding markedly. However, this effect is probably not due to loss of a direct contact between LAC9 and the T residue, since such a loss should have been detected by the filter-binding assay. The loss of binding is more likely due to a conformational change in the DNA helix induced by loss of the base (4). The ±3 and ±4 positions do not appear to be important for binding affinity or specificity as judged by missing contact footprinting and the filter-binding assay (Table 2). Bases at the ±2 or ±4 positions probably do not contribute to sequence-specific or high-affinity binding by indirect means involving DNA conformation (reviewed in reference 12). If they did, we would expect base pair changes to reduce LAC9 binding. Also, the natural UAS would be expected to show a pattern of base pairs at these positions, but they do not (Table 1). The reason for weak binding by the −3G/+3C UAS may be due to a conformational effect on DNA. It is noteworthy that none of the natural UASs (Table 1) has a −3G/+3C sequence.

The center base does not appear to make significant direct contact to LAC9 as measured by the filter-binding or missing-contact assay. Deletion of the base pair reduced binding eightfold, a value not much different than that for LAC4UASI, a UAS that is necessary for normal function of the lactose-galactose regulon (3, 13, 18). In spite of our data, the center base pair must play an essential role in LAC9 binding or some other LAC9 activity, since all natural *K. lactis* UASs contain 17, not 16, bp (Table 1). The role of the center base pair in DNA binding may be to optimize the spacing between the two halves of the asymmetric UAS, as has been noted for the DNA-binding site of GCN4 (22). Activities of LAC9, besides DNA binding, that might depend on the size and structure of the UAS include transcription activation, interaction with a negative regulator, and cooperative interactions between LAC9 molecules bound to adjacent UAS (8). Finally, the ability of LAC9(85-228*) to compensate for loss of the center base may reflect a less-rigid protein conformation than is found in the larger, native 865-amino-acid LAC9 protein. The ability of the LAC9(85-228*) dimer to bind a UAS lacking the center base indicates that the two LAC9 monomers that are contacting the DNA in adjacent major grooves must be flexible enough to adjust to the 36° change in relative orientation and the 0.34-nm decrease in the distance between the two adjacent major grooves caused by deletion of the center base pair.

Examination of the data summarized in Fig. 4 shows that LAC9 binds preferentially to one half of the UAS. Other regulatory proteins such as the repressor of phage lambda (20) bind preferentially to one half of a partially symmetric binding site. Although our data were obtained by using a UAS with symmetric half sites, we predict that LAC9 binds preferentially to one half of each natural, asymmetric UAS.

The hydroxyl radical footprint data reveal that LAC9 protects sugars at the center of each strand of the UAS from hydroxyl radical attack. These two strands form a minor groove, as shown in Fig. 4. Protected sugars are not symmetrically arranged, supporting the idea that LAC9 binds preferentially to one half of the cUAS. The protection pattern suggests that LAC9 contacts positions 6, 7, and 8 in each half of the UAS, which are separated by more than one turn of the DNA helix, and twists part way around the DNA, protecting a broad region of the minor groove between the adjacent major-groove contacts. Such a broad region of protection is unique. Further experiments will be needed to determine if LAC9 contacts phosphates in this protected region near the center of the UAS.

Previous data (17, 19, 28) showed that LAC9 and GAL4 bind the same family of UASs, and it had been argued that they have very similar DNA-binding domains (26). This hypothesis is supported also by the similarity of the hydroxyl radical footprints of LAC9 and GAL4 (Fig. 4). Slight differences in the footprints may be due to unique protein-DNA contacts or to the different UASs that were used for footprinting. Since LAC9 and GAL4 are members of a family of fungal transcription activator proteins characterized by the presence of a C6 zinc finger, we predict that other members of this family that bind 17-bp UASs contact the backbone of DNA in a manner similar to that shown for LAC9 and GAL4.

The results presented in this paper identify several bases in the UAS that are necessary for LAC9 binding. Some of these bases most likely make contacts in the major groove of the DNA with amino acid residues in LAC9. Recent results show that at least some of these LAC9 residues are located in a region of 14 amino acids adjacent to the carboxyl side of the C6 zinc finger (27). It is not clear yet if amino acids in the C6 zinc finger region also contact the UAS.

ACKNOWLEDGMENTS

Y.-D.C.H. and K.N. contributed equally to this work.

This work was supported by Public Health Service grant GM-22749 from the National Institutes of Health to R.C.D. and by Biomedical Research Support grant RR05374 from the Biomedical Research Support Branch, Division of Research Facilities and Resources, National Institutes of Health.

REFERENCES

1. **Andre, B.** 1990. The UGA3 gene regulating the GABA catabolic pathway in *Saccharomyces cerevisiae* codes for a putative zinc-finger protein acting on RNA amount. Mol. Gen. Genet. **220**:269–276.
2. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1987. Current protocols in molecular biology, p. 2.6.1. John Wiley & Sons, Inc., New York.
3. **Breunig, K. D., and P. Kuger.** 1987. Functional homology between the yeast regulatory proteins GAL4 and LAC9: LAC9-mediated transcriptional activation in *Kluyveromyces lactis* involves protein binding to a regulatory sequence homologous to the GAL4 protein-binding site. Mol. Cell. Biol. **7**:4400–4406.
4. **Brunelle, A., and R. F. Schleif.** 1987. Missing contact probing of DNA-protein interactions. Proc. Natl. Acad. Sci. USA **84**:6673–6676.
5. **Carey, M., H. Kakidani, J. Leatherwood, F. Mostashari, and M. Ptashne.** 1989. An amino-terminal fragment of GAL4 binds DNA as a dimer. J. Mol. Biol. **209**:423–432.
6. **Corton, J. C., and S. A. Johnston.** 1989. Altering the DNA-binding specificity of GAL4 requires sequences adjacent to the zinc finger. Nature (London) **340**:724–727.
7. **Dickson, R. C., and J. S. Markin.** 1980. Physiological studies of β-galactosidase induction in *Kluyveromyces lactis*. J. Bacteriol. **142**:777–785.
8. **Dickson, R. C., and M. I. Riley.** 1989. The lactose-galactose regulon of *Kluyveromyces lactis*, p. 19–40. *In* P. J. Barr, A. J. Brake, and P. Valenzuela (ed.), Yeast genetic engineering.

Butterworths, Boston.

9. **Evans, R. M., and S. M. Hollenberg.** 1988. Zinc fingers: gilt by association. Cell **52**:1–3.
10. **Garner, M. M., and A. Revzin.** 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: applications to components of the *E. coli* lactose operon regulatory system. Nucleic Acids Res. **9**:3047–3060.
11. **Halvorsen, Y.-D. C., K. Nandabalan, and R. C. Dickson.** 1990. LAC9 DNA-binding domain coordinates two zinc atoms per monomer and contacts DNA as a dimer. J. Biol. Chem. **264**: 13283–13289.
12. **Harrison, S. C., and A. K. Aggarwal.** 1990. DNA recognition by proteins with the helix-turn-helix motif. Annu. Rev. Biochem. **59**:933–969.
13. **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.
14. **Marsh, J. L., M. Erfle, and E. J. Wykes.** 1984. The pIC plasmid and phage vestors with versatile cloning sites for recombinant selection by insertional inactivation. Gene **32**:481–485.
15. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
16. **Pan, T., and J. E. Coleman.** 1990. GAL4 transcription factor is not a "zinc finger" but forms a $Zn(II)_2Cys6$ binuclear cluster. Proc. Natl. Acad. Sci. USA **87**:2077–2081.
17. **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.
18. **Ruzzi, M., K. D. Breunig, A. G. Ficca, and C. P. Hollenberg.** 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.
19. **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.
20. **Sarai, A., and Y. Takeda.** 1989. Lambda repressor recognizes the approximately 2-fold symmetric half-operator sequences asymmetrically. Proc. Natl. Acad. Sci. USA **86**:6513–6517.
21. **Seeman, N. C., J. M. Rosenberg, and A. Rich.** 1976. Sequence-specific recognition of double helical nucleic acids by proteins. Proc. Natl. Acad. Sci. USA **73**:804–806.
22. **Sellers, J. W., A. C. Vincent, and K. Struhl.** 1990. Mutations that define the optimal half-sites for binding yeast GCN4 activator protein and identify an ATF/CREB-like repressor that recognizes similar DNA sites. Mol. Cell. Biol. **10**:5077–5086.
23. **Sienbenlist, U., and W. Gilbert.** 1980. Contacts between *E. coli* RNA polymerase and an early promoter of phage T7. Proc. Natl. Acad. Sci. USA **77**:122–126.
24. **Takeda, Y., A. Sarai, and V. Rivera.** 1989. Analysis of the sequence-specific interactions between cro repressor and operator DNA by systematic base substitution experiments. Proc. Natl. Acad. Sci. USA **86**:439–443.
25. **Tullius, T. D., and B. A. Dombroski.** 1986. Hydroxyl radical "footprinting": high-resolution information about DNA-protein contacts and application to λ repressor and cro protein. Proc. Natl. Acad. Sci. USA **83**:5469–5473.
26. **Witte, M. M., and R. C. Dickson.** 1988. 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*. Mol. Cell. Biol. **8**:3726–3733.
27. **Witte, M. M., and R. C. Dickson.** 1990. The C6 zinc finger and adjacent amino acids determine DNA binding specificity and affinity in the yeast activator proteins LAC9 and PPR1. Mol. Cell. Biol. **10**:5128–5137.
28. **Wray, L. V., Jr., 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.