Protein Science (1995), *4*:1832–1843. Cambridge University Press. Printed in the USA.

UME6, a negative regulator of meiosis in *Saccharomyces cerevisiae*, contains a C-terminal Zn_2Cys_6 binuclear cluster that binds the URS1 DNA sequence in a zinc-dependent manner

STEPHEN F. ANDERSON,[1] CAMILLE M. STEBER,[2]
ROCHELLE EASTON ESPOSITO,[2] AND JOSEPH E. COLEMAN[1]

[1] Department of Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520-8114
[2] Department of Molecular Genetics and Cell Biology, The University of Chicago, Chicago, Illinois 60637

(RECEIVED May 9, 1995; ACCEPTED June 30, 1995)

Abstract

UME6 is a protein of 836 amino acids from *Saccharomyces cerevisiae* that acts as a repressor and activator of several early meiotic genes. UME6 contains, near the C-terminus, the amino acid sequence -^{771}C-X_2-C-X_6-C-X_6-C-X_2-C-X_6-C-, in which the spacings of the six Cys residues are identical to those found in 39 N-terminal Cys-rich DNA binding subdomains of fungal transcription factors. This sequence has been shown in GAL4 and other proteins to form a zinc binuclear cluster. In spite of the different location, the C-rich sequence, cloned and overproduced within the last 111 amino acid residues of UME6, UME6(111), forms a binuclear cluster and exhibits a Zn-dependent binding to the URS1 DNA sequence. The latter, TAGCCGCCGA, is required for the repression or activation of meiosis-specific genes by UME6. UME6(111) contains 1.8 ± 0.4 mol Zn/mol protein and the Zn can be exchanged for Cd to yield a protein containing 1.9 ± 0.1 mol Cd/mol protein. At 5 °C, $^{113}Cd_2$UME6(111) shows two ^{113}Cd NMR signals, with chemical shifts of 699 and 689 ppm, similar to those observed for $^{113}Cd_2$GAL4(149). The magnitude of these chemical shifts suggests that each ^{113}Cd nucleus is coordinated to four -S^- ligands, compatible with a $^{113}Cd_2$ cluster structure in which two thiolates form bridging ligands. The entire UME6 gene has been cloned and overexpressed and binds more tightly to the URS1 sequence than the zinc binuclear cluster domain alone. DNase I footprints of UME6 on URS1-containing DNA show that the protein protects the phosphodiesters of the 5′-CCGCCG-3′ region within the URS1 sequence.

Keywords: binuclear cluster; meiosis; transcriptional repressor; UME6; URS1

UME6 (CAR80/CARGR1) is a transcriptional regulator of early meiotic gene expression in *Saccharomyces cerevisiae* (Strich et al., 1989) that contains near its C-terminus the sequence -^{771}CWI^{774}CRLRKKK^{781}CTEERPH^{788}CFN^{791}CERLKLD^{798}C-, a sequence 61% homologous to that of the sequence forming a zinc binuclear cluster in the yeast transcriptional activator GAL4 (Strich et al., 1994). UME6 is one of six genes discovered during screening for yeast mutants that express early meiotic genes during vegetative growth (Strich et al., 1989). Increases in mRNA from the early meiotic genes SPO11, SPO13, and SPO16 observed in cells mutant for UME6 suggested that the gene product normally functions as a negative regulator (Strich et al., 1989). The UME6 gene has been shown to be allelic with the CAR80 gene, which codes for the repressor of the arginase gene (Park et al., 1992). Subsequent sequencing of both genes showed them to be identical (Strich et al., 1994).

Repression of the arginase gene has been shown to require the presence of the DNA sequence, 5′-TAGCCGCCGA-3′, upstream of the promoter. This sequence, known as URS1, has been shown by deletion and site-directed mutagenesis to be the site of action for UME6/CAR80 (Park et al., 1992). The URS1 sequence has been found to be required for the repression and meiotic activation of a number of genes expressed early in meiosis including SPO13 (Buckingham et al., 1990) and IME2 (Bowdish & Mitchell, 1993), as well as a number of other genes whose regulation is associated with carbon and nitrogen starvation (Sumrada & Cooper, 1987). The product of the IME2

Reprint requests to: Joseph E. Coleman, Department of Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520-8114; e-mail: coleman@zinc.csb.yale.edu.

Abbreviations: BME, β-mercaptoethanol; IPTG, isopropyl-β-D-thiogalactopyranoside; UME6, unscheduled meiotic expression gene 6; URS1, upstream repression sequence 1; PMSF, phenylmethyl sulfonyl fluoride; BSA, bovine serum albumin.

gene is a key positive regulator of early meiotic genes, and deletion of IME2 results in sporulation-deficient cells (Smith & Mitchell, 1989). This critical gene has a URS1 sequence located at positions −552 to −543 and is repressed by UME6 (Bowdish & Mitchell, 1993). In addition to functioning as a repressor, UME6 also appears to be necessary for activation of IME2 under meiotic conditions: transcription of IME2 is activated by IME1, a protein expressed only in **a**/α diploid cells under conditions of nitrogen starvation (Mitchell et al., 1990; for a review, see Mitchell [1994]). However, this activation does not occur in the absence of UME6 (Bowdish & Mitchell, 1993), suggesting that UME6 could contribute to activation of IME2 through direct or indirect interactions with IME1.

Thirty-nine different fungal transcription factors contain N-terminal DNA-binding domains exhibiting a Cys-rich amino acid sequence in which the Cys residues occupy the spacing C-X_2-C-X_6-C-X_{5-8}-C-X_2-C-X_{6-8}-C, the same spacing observed in UME6. The transcription factor, GAL4, the prototype for this group, binds to the UAS_G DNA sequence located upstream of several genes coding for the enzymes involved in galactose metabolism in *S. cerevisiae* (Bram & Kornberg, 1985; Giniger et al., 1985; for a review of galactose metabolic regulation, see Johnston [1987]). The N-terminal 62 residues of GAL4 containing the Cys-rich sequence, cloned and overexpressed in *Escherichia coli*, bind two Zn or Cd ions, which are required for DNA binding (Pan & Coleman, 1989). ^{113}Cd-^{1}H heteronuclear NMR experiments show that the six conserved cysteines form the ligands for a Cd_2Cys_6 binuclear cluster in which two of the cysteine S^- bridge the two metal ions (Pan & Coleman, 1990a, 1990b; Gadhavi et al., 1991; Gardner et al., 1991). The crystal structures of the N-terminal 65-residue fragment of Cd_2GAL4 complexed with the UAS_G DNA and the structure of the DNA-binding domain of another zinc cluster protein, Zn_2PPR1, complexed with DNA have been solved (Marmorstein et al., 1992; Marmorstein & Harrison, 1994). NMR solution structures of Cd_2GAL4(65) (Baleja et al., 1992), Zn_2GAL4(43) (Kraulis et al., 1992), Zn_2GAL4(62) (Shirakawa et al., 1993), and a third member of the class, LAC9(61) (Gardner et al., 1995), all confirm the general fold of the zinc binuclear cluster domain.

In this paper, we describe the cloning and overproduction of the complete UME6 protein, UME6(838), as well as a C-terminal 111-residue fragment, UME6(111), containing the Cys-rich sequence homologous to that forming binuclear zinc clusters in the transcription factors. We show that the C-terminal subdomain binds two Zn or Cd ions and the ^{113}Cd-NMR spectrum of the ^{113}Cd-substituted derivative is consistent with the formation of a Cd binuclear cluster. We also show that UME6(111) and UME6(838) specifically bind the URS1 sequences from the SPO13 and IME2 genes. UME6(111) does so in a zinc-dependent manner.

Results

Metal content and ^{113}Cd-NMR of UME6(111)

Cloned, purified UME6(111) contains 1.8 ± 0.4 mol of Zn/mol protein as determined by atomic absorption spectroscopy. The ^{113}Cd-derivative prepared by metal ion exchange contains 1.9 ± 0.1 mol of Cd/mol protein. At 35 °C, the ^{113}Cd NMR spectrum of $^{113}Cd_2$UME6(111) shows a single ^{113}Cd resonance 695 ppm downfield of the signal for $^{113}Cd(ClO_4)_2$ (Fig. 1A). An estimate of the amount of ^{113}Cd represented by this signal was made by integration of the area under the ^{113}Cd NMR signal and comparison to a similar integration of the resonance for a $^{113}Cd(ClO_4)_2$ standard of known concentration, compensating for the difference in relaxation times. The 695-ppm resonance represents approximately one ^{113}Cd nucleus/mol of protein.

At 5 °C, a second ^{113}Cd resonance appears at 687 ppm with an area only slightly less than that of the original signal (Fig. 1B). Thus, the NMR signal from one of the two bound ^{113}Cd ions in the protein is broadened beyond detection by a chemical exchange process at temperatures above 5 °C. The ^{113}Cd resonance observed at 35 °C for $^{113}Cd_2$UME6(111) is broader, $v_{1/2}$ = 163 Hz, than the resonances of $^{113}Cd_2$LAC9(61) at 35 °C, $v_{1/2}$ = 109 Hz and 122 Hz for the downfield and upfield signals, respectively (Fig. 2C). On the other hand, the ^{113}Cd resonances of the larger LAC9(144) at 705 and 692 ppm have $v_{1/2}$ values of 178 and 222 Hz, respectively, at 25 °C (Pan et al., 1990). The line width of the ^{113}Cd resonance at 695 ppm in UME6(111) is consistent with a protein of a size intermediate between that of LAC9(61) and LAC9(144), and therefore does not seem to be significantly modulated by the exchange process that broadens the other ^{113}Cd resonance (see Discussion).

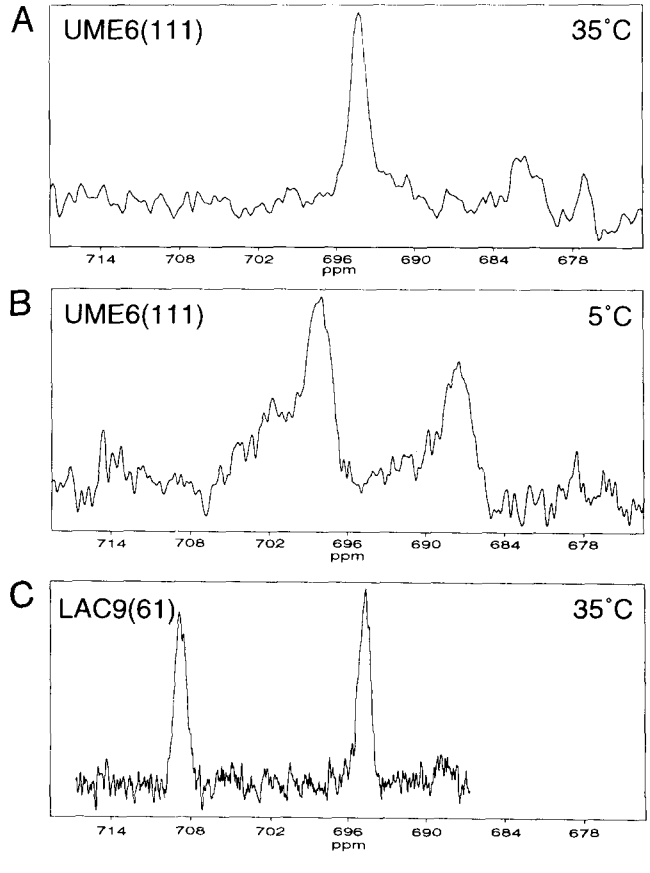

Fig. 1. Cadmium-113 NMR spectra of 0.1 mM UME6(111) in 10 mM KP_i, pH 7.5, 300 mM NaCl at (**A**) 35 °C and (**B**) 5 °C. Spectrum A required 750,000 scans, whereas the line broadening at the lower temperature of spectrum B required 4,960,000 scans. **C**: Spectrum of Cd_2LAC9(61) at 35 °C plotted to the same scale for comparison.

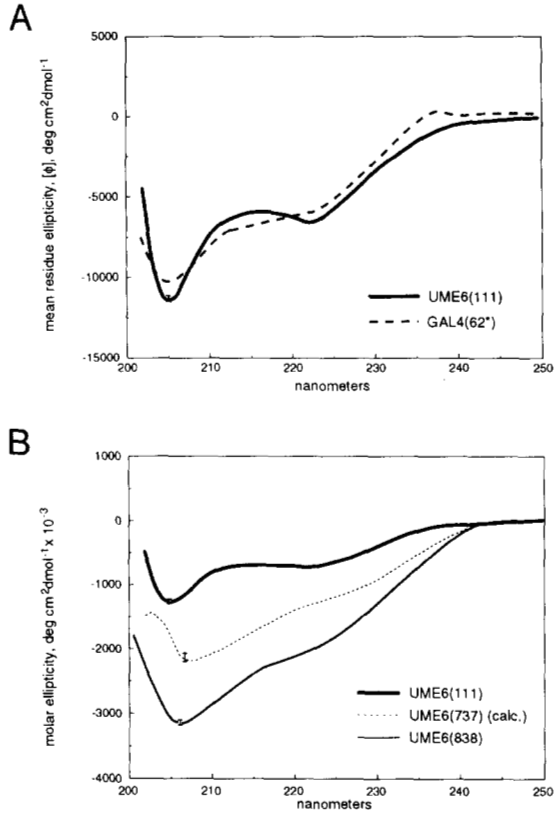

Fig. 2. A: CD spectrum of UME6(111) plotted as mean residue ellipticity (solid line). CD spectrum of GAL4(62*) is included for comparison (dashed line). **B:** CD spectra of UME6(838) (thin solid line), UME6(111) (thick solid line), and the difference spectrum between UME6(838) and UME6(111) (dotted line) representing the contribution of the N-terminal 727 residues, plotted as molar ellipticity. Error bars indicate the magnitude of the noise level within the spectra. Conditions were 1 μM for UME6(838) and 100 μM for UME6(111), both in 10 mM KP_i, 300 mM NaCl, pH 7.5, 25 °C.

Conformation of UME6(111) and UME6(838) as inferred from CD

The UV CD spectrum of Zn_2UME6(111) (heavy line, Fig. 2A) is similar in shape and magnitude of mean residue ellipticity to that observed for the 62-residue N-terminal fragment of GAL4, Zn_2GAL4(62*) (Pan & Coleman [1990a], dashed line in Fig. 2A). In GAL4(62) and LAC9(61), the zinc binuclear cluster involves approximately 40 residues, or two-thirds of the fragment. An estimate of the α-helical secondary structure present in UME6(111), calculated by the method of Greenfield and Fasman (1969), yields approximately 17% helix. This value is consistent with the approximately 12 residues of helix found within the cluster region of GAL4(62) (Baleja et al., 1992; Kraulis et al., 1992; Marmorstein et al., 1992; Shirakawa et al., 1993).

The molar ellipticity of UME6(838) in 0.1% Tween20 is compared to that of UME6(111) in Figure 2B. The contribution to the molar ellipticity of the whole protein by the C-terminal 111 residues, assuming its structure remains the same in the context of the whole protein, represents about 40% of the total signal (thick line, Fig. 2B). The ellipticity associated with the peptide bond chromophores of UME6(838), other than those of the C-terminal 111 residues, is thus rather small, with a mean residue ellipticity of only 3,000 deg cm^2 dmol^{-1} at 208 nm (dotted line, Fig. 2B). The overall profile of the difference CD spectrum is similar to that determined for proteins having only coil conformation (Saxena & Wetlaufer, 1971).

Analysis of the hydrophilicity of UME6(836)

UME6 has an unusual amino acid composition compared to the average globular protein in that the ratio of hydrophilic to hydrophobic residues is 2.02, more than twice the ratio of 0.91 found in the average globular protein. GAL4(881), in comparison, has a more normal ratio of 1.24. The hydrophobicity of UME6(836) can be analyzed using the computer program SEQSEE (Wishart et al., 1994). The output of this program is a running average of the hydrophobicity over a moving, 12-residue, nonweighted averaging "window" calculated using the Kyte-Doolittle hydrophobicity scale. This scale assigns a value of +4.5 to polyisoleucine and −4.5 to polyarginine (Kyte & Doolittle, 1982). Using an averaging window of this size, stretches of residues where the average hydrophobicity is between 1 and 2 correlate with regions of from five to seven sequential highly hydrophobic residues. Analysis of proteins with known tertiary structures show that when the average hydrophobicity over a window of this size is between 1 and 2, the residues are generally found to be part of the hydrophobic core.

In Figure 3, the hydrophobicity plots of UME6(836) and GAL4(881) are compared. The peaks in the plot for GAL4 are numerous and hydrophobic enough that they are likely to represent regions of the polypeptide involved in the formation of

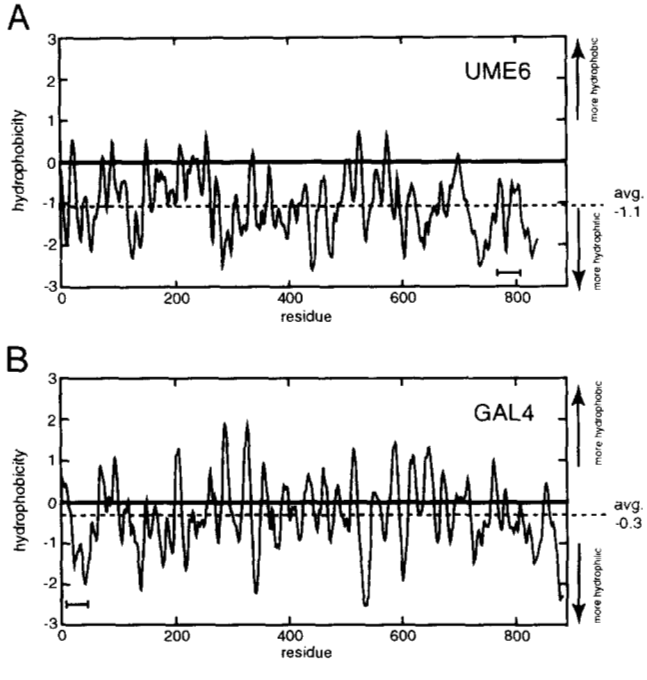

Fig. 3. Hydrophobicity plots of (**A**) UME6 and (**B**) GAL4. Positive values indicate greater hydrophobicity, as calculated from the Kyte-Doolittle hydrophobicity scale (see text). A value of 4.5 corresponds to polyisoleucine, whereas a value of −4.5 corresponds to polyarginine. An unweighted 12-residue window was used to calculate a running average and smooth the data. Locations of the binuclear cluster DNA-binding domains are indicated by crossbars below each line.

a hydrophobic core (Fig. 3B), thus GAL4(881) is predicted to have a defined tertiary structure in solution. In contrast, the plot for UME6(836) shows that the most hydrophobic regions of UME6(836) are far less hydrophobic than the corresponding regions of GAL4 or other proteins known to fold (Fig. 3A). UME6 appears to lack sufficient stretches of hydrophobic residues to form a significant hydrophobic core, suggesting that a significant amount of UME6 may be disordered in solution. The same conclusion is reached by a more empirical analysis based on the average hydrophobicity: the hydrophobicity averaged over all 836 residues of UME6 is −1.06, whereas GAL4 has an average hydrophobicity of −0.30, close to the average value of −0.26 for globular proteins. No protein with an average hydrophobicity lower than −0.7 is known to fold (D. Wishart, pers. comm.). Thus, on the empirical basis of its low average hydrophobicity alone, extensive folding of the UME6(836) polypeptide in aqueous solution is unlikely, compatible with the CD spectrum (thin line, Fig. 2B).

One factor not included in a Kyte-Doolittle hydrophobicity analysis is the possibility of electrostatic barriers to protein folding. Although UME6 contains a comparable percentage of charged amino acid residues to GAL4 (22-19%, respectively), the net charge from the 45 aspartate, 40 glutamate, 47 lysine, and 36 arginine residues found in GAL4(881) is −2, whereas UME6(836) contains 48 Asp, 31 Glu, 62 Lys, and 46 Arg residues, for a net charge of +29. In UME6, this charge is spread out over a protein containing 45 fewer residues than GAL4, leading to a significantly increased charge density. Although it is not possible to say exactly what effect this additional charge might have on the folding of UME6, the resulting electrostatic repulsion could result in the UME6 polypeptide taking up an extended conformation in solution.

Gel retardation of URS1-containing DNA by UME6(111) and UME6(838)

A gel retardation titration of the 22-bp oligonucleotide containing the base sequence of the URS1 from the SPO13 gene with UME6(111) is shown in Figure 4. The protein concentration was varied from 0.01 to 50 μM. In the range of protein concentration from 0.01 to 10 μM, a single major complex is formed that retains most of the DNA when the protein reaches a concentration of 10 μM. At 50 μM UME6(111) nonspecific binding or formation of protein oligomers occurs and the retarded band becomes slow moving and heterogeneous. Because the N- and C-terminal regions of UME6(111), rich in K and R residues, are susceptible to *E. coli* proteases, there is a minor faster-moving band that represents the complex with a smaller proteolytic fragment of UME6(111).

The gel retardation data were fit to the binding equation $f_b = P^n/(P^n + K_d^n)$, in which f_b is the fraction of DNA bound, P is the protein concentration, K_d is the dissociation constant, and n is the cooperativity parameter. The above data yield a K_d of 0.62 ± 0.08 μM and a cooperativity parameter of 1 (Table 1). The titration shown in Figure 4 is in the absence of competitor DNA, but the addition of a 10^4 molar excess of nonspecific DNA has relatively little effect on the binding curve.

A gel retardation of the same URS1 fragment with 500 nM Zn_2UME6(111) and 500 nM apoUME6(111) is shown in Figure 5. Both of these samples contained a 10^4 molar excess of nonspecific DNA. The UME6(111) sample used in this experiment has much more of the smaller proteolytic product present, thus the faster moving complex formed with the proteolytic fragment is more prominent than the slower moving complexes formed with more intact UME6(111) fragments. Formation of all complexes is abolished by removal of the zinc from the protein with EDTA.

A plasmid containing the DNA sequence of the URS1 site from upstream of the IME2 gene was constructed for use in gel retardation experiments as well. Cleavage of this plasmid with *Xho* I results in three DNA fragments, the URS1-containing fragment (169 bp), a fragment (93 bp) containing no sequences related to URS1 ("nonspecific"), and a third fragment, ("vector," 2.9 kb), which contains two sequences, each of which differ from URS1 by only two bases. A gel retardation titration showed that UME6(111) binds with specificity to the URS1-containing fragment and with some specificity to the "vector" fragment containing the two "pseudo-URS1" sites. The K_d val-

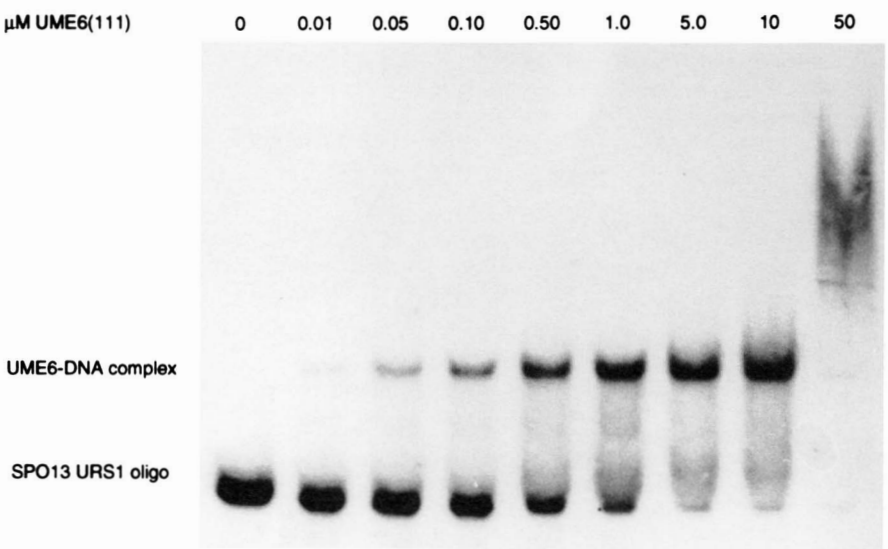

Fig. 4. Gel retardation titration of a synthetic 22-bp DNA fragment containing the SPO13 URS1 sequence with increasing amounts of UME6(111). All binding reactions were incubated 5 min in 20 mM NaP_i, 115 mM NaCl, 10 μM $ZnCl_2$, 1 mg/mL BSA, 20% glycerol before loading on a 4% polyacrylamide TB gel.

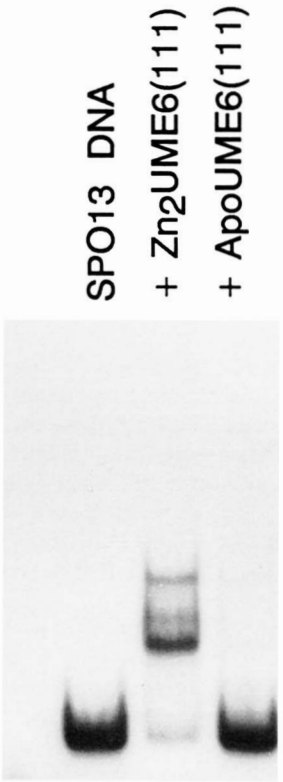

Fig. 5. Zinc-dependent binding of UME6(111) to the SPO13 URS1. Lane 1, 1 nM SPO13 DNA only; lane 2, SPO13 DNA + 500 nM Zn_2UME6 (111); lane 3, SPO13 DNA + 500 nM ApoUME6(111). Binding reactions took place in 20 mM NaP_i, 25 mM NaCl, 10 μM $ZnCl_2$, 20% glycerol, and a 10^4 excess of unlabeled salmon sperm DNA competitor.

ues for UME6(111) bound to all three of these fragments are given in Table 1.

When the full-length protein, UME6(838), is bound to even the 22-bp duplex DNA, most of the complexes formed remain in the well, although complexes corresponding to proteolytic fragments of UME6(838) can be observed entering the gel at 10 nM UME6(838) (Fig. 6). A gel retardation titration of the DNA fragments containing the URS1 sequence from the IME2 gene (169 bp) and the fragment containing only unrelated sequences (93 bp) is shown in Figure 6. Although 50% of the URS1 fragment is retained at 6 nM UME6(838), by 50 nM protein both fragments are completely retained in the well. However, the addition of 500 mM NaCl dissociates the protein from the nonspecific fragment, whereas the URS1-containing fragment is still largely retained at the high salt concentration (right lane, Fig. 6).

DNase I footprinting of UME6(838) and UME6(111)

The footprints of the C-terminal fragment, UME6(111), and the full-length protein, UME6(838), on the URS1 sequence from the SPO13 gene were determined by DNase I digestion. A densitometer tracing of the footprint of UME6(838) on both strands of the 22-bp DNA containing the URS1 sequence from the SPO13 gene is shown in Figure 7. Footprints of UME6(111) on this URS1 sequence and of both UME6(111) and UME6(838) on the URS1 sequence from the IME2 gene give similar results (Fig. 8A). A projection of the diester protections observed on B-form DNA is shown in Figure 8B. The URS1 sequence covers the 10 bp from the AT at the top of Figure 8B to the first AT after the GC-rich center of the sequence. The protection against DNase I cleavage provided by the bound UME6 is asymmetrical and concentrated in the region of the two CCG triplets of the URS1 sequence (Fig. 8B). Within the 5′-CCGCCG-3′ strand, the diester bonds 3′ to all three nucleotide residues of the first triplet are strongly protected (Fig. 7A). There is moderate protection of the bond 3′ to the first C of the second triplet, whereas the bond 3′ to the middle C of the second triplet again is strongly protected (Fig. 7A).

On the opposite strand, 5′-CGGCGG-3′, the bonds 3′ to the first C and the middle G of the first triplet are moderately protected, whereas the bond 3′ to the middle G of the second triplet is also moderately protected (Fig. 7B). On this strand, there is, in addition, a very strong protection of the bond 3′ to the C immediately 3′ to the second CGG triplet. This GC base pair is still considered part of the URS1 sequence but is not so strongly conserved among the many URS1 sequences as the two CCG triplets. There are also several diester bonds at the fringes of the

Table 1. *Dissociation constants and Hill coefficients for UME6 binding to DNA*[a]

	UME6(111) K_d (nM)	UME6(111) Hill coefficient	UME6(838) K_d (nM)	UME6(838) Hill coefficient
SPO13 URS1	620 ± 80	0.62	12 ± 2	1.3
IME2 URS1	110 ± 20	1.1	5.9 ± 0.6	1.5
"Vector" 2.9 kb	210 ± 20	0.99	~7.5	N.D.
Nonspecific	~2,000	N.D.	N.D.	N.D.

[a] Values were determined by computer quasi-Newton fitting of data points to the equation: $f_b = P^n/(P^n + K_d^n)$, where f_b is the fraction of DNA bound, P is the protein concentration, n is the Hill coefficient, and K_d is the dissociation constant for the binding reaction. Error values are standard deviations as determined by the curve-fitting program. In cases where the computer did not converge on a solution due to insufficient data points, the K_d was estimated and is indicated by the ~ symbol. The row marked "Vector" 2.9 kb contains values for UME6 binding to the pBLUESCRIPT vector fragment of pI2U1, which contains two sequences that are two-base deviations from a canonical URS1 site. N.D. means not determined, because the computer could not assign a value to n based on the distribution of data points.

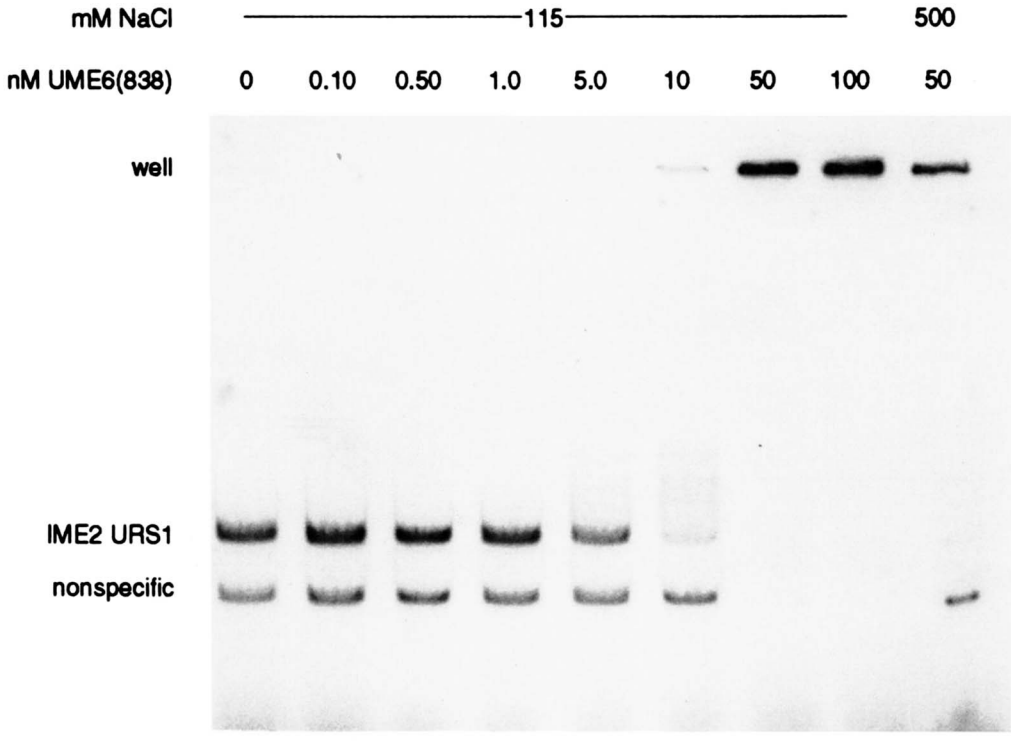

Fig. 6. Gel retardation titration of a 169-bp *Xho* I fragment containing the IME2 URS1 sequence and a 93-bp fragment with no homology to URS1. All incubations were for 5 min at 20 °C in 20 mM NaP_i, 115 mM NaCl, 10 μM $ZnCl_2$, 1 mg/mL BSA, 20% glycerol, except lane 9, in which the NaCl concentration was raised to 500 mM. Gels were 4% polyacrylamide run in 90 mM Tris-borate, pH 9.0.

URS1 site that are slightly protected from nuclease digestion by UME6 binding.

Discussion

The chemical shifts of the two ^{113}Cd NMR signals of $^{113}Cd_2$UME6(111) at 5 °C (699 and 689 ppm) are within the range of chemical shifts found for ^{113}Cd ions coordinated to four -S^- ligands, based on the correlation of chemical shift with type and number of donor atoms derived from ^{113}Cd-substituted protein sites of known donor composition (Coleman, 1993). Similar ^{113}Cd NMR resonances are observed for the transcription factors $^{113}Cd_2$GAL4(149) (707 and 669 ppm) at 35 °C (Pan & Coleman, 1990a, 1990b), and $^{113}Cd_2$LAC9(144) (706 and 691 ppm) at 25 °C (Pan et al., 1990). ^{113}Cd-^{1}H heteronuclear NMR techniques applied to both GAL4(62*) and LAC9(61) have shown that each ^{113}Cd ion in these proteins has four cysteinyl -S^- ligands, two of which are bridging ligands coordinated to both Cd ions (Gardner et al., 1991; Gardner & Coleman, 1994). This ligand arrangement has been confirmed by the crystal structure of the Cd_2GAL4(65) complex with the UAS_G DNA (Marmorstein et al., 1992) and the co-complex of Zn_2PPR1 with DNA (Marmorstein & Harrison, 1994). Therefore, the best explanation for the chemical shifts of the ^{113}Cd NMR signals of $^{113}Cd_2$UME6(111) is that the six cysteinyl residues coordinate the two Cd ions by forming a binuclear metal cluster in which two of the -S^- ligands bridge the two metal ions.

Not only is the spacing of the six cysteinyl residues in UME6(111) identical to that found in the zinc-cluster protein GAL4, but the UME6 sequence conserves the prolyl residue located two residues N-terminal to the fourth Cys residue in all the zinc-cluster transcription factors. This residue is in the *cis* conformation in the GAL4, PPR1, and LAC9 structures (Marmorstein et al., 1992; Marmorstein & Harrison, 1994; Gardner et al., 1995). The *cis* configuration appears to be necessary for the formation of a loop of six residues that places the two -C-X_2-C-X_6-C- sequences in position to form the opposite halves of the binuclear metal cluster. The sequence of UME6 also contains the basic region between the second and third cysteine ligands, which in the two crystal structures forms a short helix responsible for the sequence-specific contacts with the DNA. The CD spectrum of UME6(111) indicates sufficient helical structure to account for this recognition helix as well as the helix that forms the second half of the binuclear cluster. The Cys-rich sequence of UME6 therefore contains all the structural elements required to form the zinc binuclear cluster motif.

Although exchange broadening of ^{113}Cd resonances has been described for several ^{113}Cd-substituted Zn proteins (Coleman, 1993), the room temperature exchange broadening of the signal from one of the two ^{113}Cd ions in UME6(111) is the first -S_4 site where exchange broadening is sufficient to render the resonance undetectable at room temperature. The exchange process must be a conformational flux of the protein transmitted to the ^{113}Cd environment rather than metal ion dissociation, because the stoichiometry of the complex remains 2 Cd/mol of protein after a 24-h dialysis at 25 °C. If a dissociated metal ion species were the second species in a chemical exchange with the

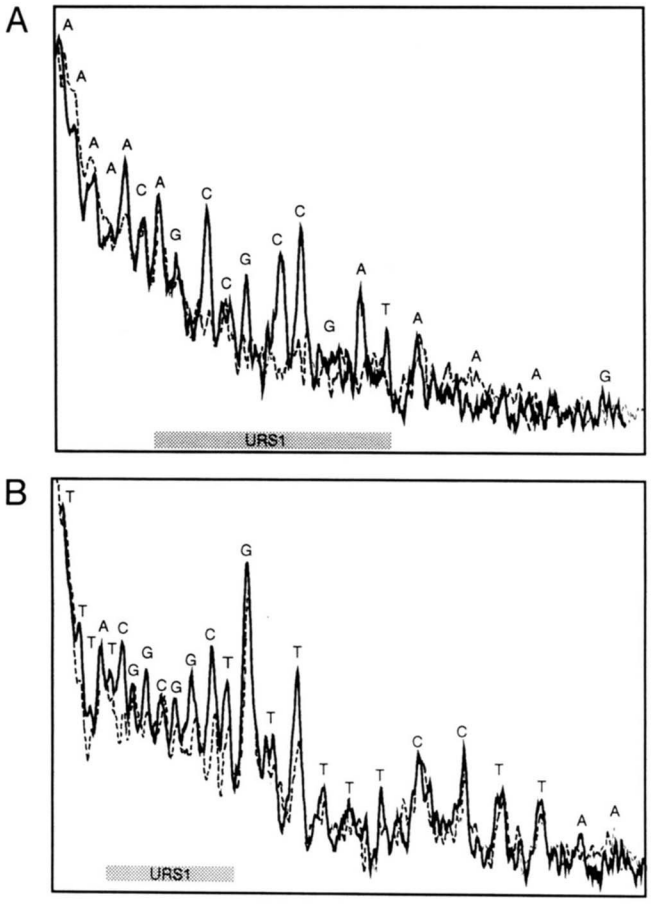

Fig. 7. Footprinting of UME6(838) on the SPO13 URS1. Densitometric traces of DNase I cleavage ladders for (**A**) 22-mer and (**B**) 30-mer strands of the SPO13 URS1 duplex in the absence (solid lines) and presence (dotted lines) of 0.1 μM UME6(838) protein. Position of the URS1 sequence in each strand is indicated by a gray box below the trace.

protein-bound ^{113}Cd, then its population would have to be at least 10% of the total in order to significantly broaden the signal of the bound ^{113}Cd (Coleman et al., 1979). Such an equilibrium would be expected to result in significant dissociation of the metal ion on even a short dialysis. However, a partial rearrangement of the ligands, e.g., exchange of a water molecule for one of the sulfur ligands, is a possible source of a second conformation of the ^{113}Cd site giving rise to the upfield resonance.

In addition to the intrinsic frequency of the conformational flux, dependent on temperature, the other variable that determines intermediate chemical exchange on the NMR time scale is the frequency difference, $\Delta\nu$, in Hz between the resonances of the two exchanging species. Thus, the temperature, the inherent chemical shift difference between the ^{113}Cd resonances characterizing the two conformations, and the field strength are the three variables that determine the intermediate exchange condition. Changes in field strength, although not changing the the chemcial shift difference, do change the $\Delta\nu$. The effects on ^{113}Cd NMR signals of changes in these three variables for the resonances of two species in chemical exchange are modeled in Coleman et al. (1979). The striking exchange modulation at 35 °C appears to differentiate the metal cluster of UME6 from those of GAL4 and LAC9. However, as noted above the conditions of frequency and temperature at which a conformational modulation will be observed to broaden the NMR signal from a ^{113}Cd complex depend on precise chemical shift differences between the two or more ^{113}Cd resonances involved (Coleman et al., 1979). The chemical shift difference and hence $\Delta\nu$ can vary significantly even for otherwise closely related ^{113}Cd sites.

If the ^{113}Cd spectrum of $^{113}Cd_2$GAL4(62) is obtained at 44 MHz, an exchange modulation of the upfield ^{113}Cd signal is observed (Pan & Coleman, 1990a). Likewise, if the temperature of acquisition for the ^{113}Cd NMR spectrum of $^{113}Cd_2$LAC9(61) is 5 °C, an exchange modulation of the upfield ^{113}Cd signal is observed (K.H. Gardner & J.E. Coleman, in prep.). For both GAL4(62) and LAC9(61), heteronuclear correlation methods have shown that the upfield ^{113}Cd NMR signals, subject to exchange modulation, arise from the ^{113}Cd ion coordinated to the more N-terminal position in the cluster, i.e., primarily to Cys residues 11, 14, and 21 in GAL4 and Cys residues 95, 98, and 105 in LAC9 (Gardner et al., 1992; Gardner & Coleman, 1994). This is the metal ion most closely associated with the short helix contributing the base-specific contacts to the DNA (Marmorstein et al., 1992). It is also the metal ion most easily exchanged in GAL4(62) (Pan & Coleman, 1990a). Thus, the exchange modulation of the upfield ^{113}Cd resonance of UME6 appears to identify this signal as coming from the ^{113}Cd ion coordinated to the more N-terminal Cys residues associated with the DNA-recognition helix.

The 12–14 residues expected to be in a helical conformation in a binuclear zinc cluster account for most of the CD spectrum observed for UME6(111). Thus, it is possible that the positively charged regions of the polypeptide flanking the cluster have little organized secondary structure. The unusually high charge density and unusually low hydrophobicity of UME6(836) predict that it is disordered in solution. This is consistent with the lack of secondary structure implied by the CD spectrum of the full-length protein. It is therefore likely that, with the exception of the binuclear cluster domain, there is little or no additional secondary structure in the UME6 gene product. The hydrophobicity analysis of UME6(836) (Fig. 3A) suggests that there is little hydrophobic core for UME6(836) to fold upon. This is in contrast to GAL4, another zinc cluster-containing protein, which contains significant organized secondary structure outside the cluster subdomain and which contains clusters of hydrophobic residues (Fig. 3B). GAL4 and most other yeast transcription factors containing the binuclear zinc cluster in their N-terminal domains have a leucine-rich sequence following the cluster that forms a dimerization domain. The resulting coiled-coil contributes prominently to the CD of the fragments of GAL4 larger than 65 residues (Pan & Coleman, 1989). No region of sequence like that of the dimerization domain of the transcription factors is found anywhere in the polypeptide of UME6. Neither does the CD spectrum of full-length protein indicate a dimerization helix. In fact, search of the amino acid sequence of UME6 and the sequences of a number of yeast zinc-cluster transcription factors for regions of homology between UME6 and the transcription factors using the computer program SEQSEE[2] found little or no homology outside of the sequence surrounding the binuclear cluster.

If folded structures do exist in UME6, they may be small domains connected by flexible linkers in a "beads-on-a-string" arrangement. Such a structure could allow regulatory regions of

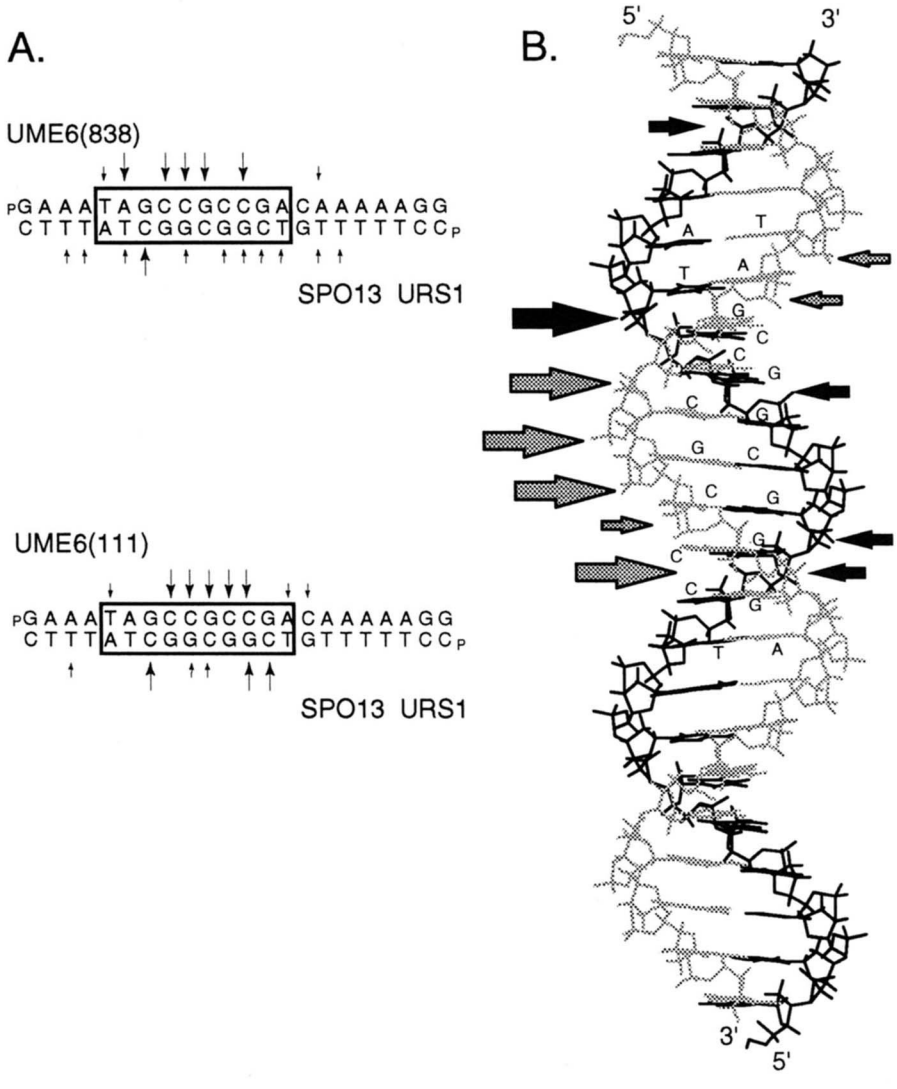

Fig. 8. Summary of UME6 footprinting data. **A:** Schematic representation of DNase I protection. Bases protected more than 50% are indicated with large arrows and bases protected less than 50% are indicated with small arrows. URS1 sequence within each DNA fragment is boxed. **B:** Consensus footprint mapped to B-form DNA. Backbone positions protected strongly in both footprints are drawn as large arrows and backbone positions protected strongly in one or weakly in both footprints are drawn as small arrows. Black arrows point to protected phosphodiester bonds on the black strand, and gray arrows point to protected phosphodiester bonds on the gray strand. Bases within the URS1 site are labeled.

UME6 to act on the transcriptional complex at distances considerably further from the URS1 binding site than possible with a compact globular protein. It is possible that UME6 interacts in vivo with several other proteins of the transcription complex, e.g., IME1, which could control much of the folding of UME6. IME1, the transcriptional activator has been reported to require UME6 for activation of the IME2 gene (Bowdish & Mitchell, 1993) and is one of the few proteins that we have found to have a hydrophobicity as low as that of UME6. If IME1 and UME6 associate directly or indirectly through a third protein, formation of a common hydrophobic core might be a simple means of forming a heterodimer. IME1 has a net charge of −17 and if it associates with UME6 it is possible that neutralization of some of the 29 net positive charges of UME6 would allow additional domains of both polypeptides to fold.

Several observations associated with the regulation of gene expression by UME6 could be explained by a flexible overall structure. URS1 sites have been found both upstream and downstream of promoters for genes regulated by UME6. A long flexible "tether" for UME6 could allow regulatory regions of UME6 to contact the transcription complex without requiring a rigid positioning of the URS1 binding site. UME6 regulation is independent of the orientation of the URS1 DNA sequence relative to the promoter being regulated (Buckingham et al., 1990). In the case of many transcription factors, this orientation independence derives from the formation of symmetrical dimers of the proteins on palindromic DNA sequences. URS1 is not palindromic and there is no evidence from the gel retardation studies or the footprints that UME6(111) or UME6(838) form dimers on the DNA. The shorter N-terminal fragments of GAL4, GAL4(62), and GAL4(63), containing only the binuclear cluster, do not dimerize significantly in solution. Although they recognize the UAS_G DNA specifically, they show relatively weak binding, $K_d = 65\ \mu M$, and the binding steps for the individual monomers can be detected by gel retardation titrations (Rodgers & Coleman, 1994). Although GAL4(62) and GAL4(65) dimerize on the DNA (Marmorstein et al., 1992), the larger fragment, GAL4(149), dimerizes in solution and the K_d drops to 13 nM (Carey et al., 1989; Rodgers & Coleman, 1994). The entropic advantage associated with dimer formation in solution accounts for a large part of the increase in affinity for the palindromic UAS_G site.

UME6(111) has a considerably higher affinity for the URS1 sequence than the short fragments of GAL4 have for the UAS_G

sequence, yet both the footprint and the cooperativity parameter of 1 in the gel retardation titrations suggest that UME6(111) binds the DNA as a monomer. In contrast, detailed gel retardation titrations with N-terminal fragments of GAL4 >100 residues show cooperativity parameters of 2 (Rodgers & Coleman, 1994). The high positive charge density surrounding the cluster of UME6(111) (10 Lys and Arg residues in the N-terminus and 15 Lys and Arg residues in the C-terminus) may enhance the DNA binding of the 111-residue UME6 fragment relative to the binding of GAL4(62*) through nonspecific interactions with the phosphodiester backbone. These basic flanking regions do not bind DNA in the absence of a folded binuclear cluster domain, because removal of the metal from the cluster by dialysis destroys UME6(111) binding to the URS1 (Fig. 5). The major DNA-binding structure in UME6 must therefore be the binuclear cluster. The similarities of the footprints suggest that both UME6(838) and UME6(111) recognize the DNA with essentially the same DNA-binding surface. The 50-fold higher affinity of UME6(838) for the URS1 DNA must therefore arise from subtle changes in the DNA-binding interactions, which remain dominated by the contribution of the binuclear cluster.

The footprints of UME6(111) and UME6(838) suggest that UME6 contacts the DNA primarily at the 5′-CCGCCG-3′ sequence. The binding of zinc binuclear cluster-containing proteins to GC-rich sequences, specifically CCG or CGG triplets, has been a characteristic of transcription factors containing this zinc-binding motif (de Rijcke et al., 1992; Dhawale & Lane, 1993 and references therein), although in this case the binding site is a direct repeat of the CCG sequence without a spacer, an arrangement not encountered previously. The pattern of DNase I protection, primarily on one strand of the double helix, is compatible with the binding of a single zinc cluster domain with one helix entering the major groove. The occlusion of the DNA backbone of one strand more than the other is also observed in the space-filling representation of the GAL4(65)-UAS_G DNA structure (Marmorstein et al., 1992). The cluster of UME6 differs from those of the transcriptional activators of the GAL4 class in that it is flanked by highly positively charged peptide sequences. These may contact the phosphodiester backbone and be responsible for some of the protection of the DNA outside of the CCG triplets in the UME6(838) footprint.

Prior to the discovery of a zinc binuclear cluster motif in UME6, this zinc-binding motif was confined to the N-terminal sequences of 39 yeast transcription factors, a fairly homologous group of proteins. The finding of the zinc binuclear cluster in the C-terminus of another class of protein that bears little homology to the earlier transcriptional activators suggests that, like several of the 10 or more zinc-binding motifs now known in various proteins involved in gene expression, the zinc cluster motif may have been inserted in a variety of DNA-binding proteins of widely varying structure and function.

Materials and methods

Cloning of UME6(111)

A fragment of the UME6 gene starting with the triplet for Ala^{727} and extending downstream of the translational stop signal was amplified by PCR from the plasmid pPL5915 (Strich et al., 1994). The upstream primer changes the codon for Lys^{726} to an ATG and also introduces an *Nde* I site just upstream of this ATG. The downstream primer anneals to the DNA 384–403 bp downstream of the stop codon. Because there is a second *Nde* I site 264 bp downstream of the TAA, digestion of the amplified fragment with *Nde* I yields a 601-bp *Nde* I fragment. The latter was ligated into the *Nde* I site of pAR3039 (Rosenberg et al., 1987) to form pUME6(111). pUME6(111) was electroporated into *E. coli* BL21(DE3) (Studier & Moffatt, 1986) with a Bio-Rad Gene Pulser using the conditions of Dower et al. (1988). The transformed cells were grown in LB media, 37 °C, to an OD_{600nm} of 0.7, and induced with 0.3 g/L IPTG (Boehringer-Mannheim). Maximal induction is attained after 4 h.

The amino acid sequence of UME6(111) is $MAKS^{730}KAKQ$ $SSKKRP^{740}NNTTSKSKAN^{750}NSQESNNATS^{760}STSQGTRS$ $RT^{770}GCWICRLRKK^{780}KCTEERPHCF^{790}NCERLKLDCH$ $^{800}YDAFKPDFVS^{810}DPKKKQMKLE^{820}EIKKKTKEAK^{830}R$ RAMKKK.

Purification of UME6(111)

Seven grams of cells induced for UME6(111) were resuspended in 35 mL of 20 mM Tris-HCl, pH 7.5, 300 mM NaCl, 10 mM BME, 1 mM $ZnCl_2$, sonicated twice for 5 min, and centrifuged at 27,000 × *g* for 20 min. The supernatant was diluted fivefold in buffer A (20 mM Tris-HCl, pH 7.5, 75 mM NaCl, 1 mM BME, 500 μM $ZnSO_4$) and loaded onto a Trisacryl-SP cation-exchange column (IBF Biotechnics, Columbia, Maryland). The column was washed with three volumes of buffer A and eluted with a gradient of 75–2,000 mM NaCl in buffer A. Fractions containing UME6(111) were pooled, diluted 10-fold with buffer A, and loaded onto an ssDNA-cellulose column. The column was washed with three volumes of buffer A and eluted with a gradient of 75–1,000 mM NaCl. The fractions containing UME6(111) were diluted twofold with buffer A and loaded onto an Affi-Gel Blue column (Bio-Rad). The column was washed with three volumes of buffer A and eluted with a gradient of 500–2,000 mM NaCl. Fractions containing pure UME6(111) were pooled and concentrated by ultrafiltration on an Amicon YM-3 membrane. The UME6(111) was over 95% pure as judged by SDS-acrylamide gel electrophoresis using silver staining of the gel.

Preparation of ApoUME6(111)

Protein, diluted to 20 μM in 10 mM K-acetate, pH 5.0, 10 mM EDTA, 500 mM NaCl, 1 mM BME, was dialyzed (2,000 MW cutoff dialysis tubing), against 6 L of the same buffer for 48 h at 4 °C with one change of buffer. The protein was then dialyzed against 10 mM Tris, pH 7.5, 500 mM NaCl, 1 mM BME for 24 h, and concentrated to 1 mM using an Amicon stirred cell with a YM-3 membrane.

Preparation of $^{113}Cd_2$UME6(111)

Zn_2UME6(111) in 50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 1 mM BME, was treated under nitrogen with a threefold molar excess of $^{113}CdCl_2$ (prepared from 95.3 atom% excess ^{113}Cd metal; Prochem/U.S. Services, Summit, New Jersey) for 36 h at 25 °C. Excess ^{113}Cd was removed by dialysis against 20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 1 mM BME, at 4 °C for 36 h, followed by dialysis against 20 mM K-phosphate, pH 5.75, 500 mM NaCl, for 36 h and concentration to 0.5 mL by ultrafiltration. Protein

concentrations were determined from the OD_{280nm} of UME6(111) dilutions in 6 M guanidinium-HCl, 20 mM K-phosphate, employing an extinction coefficient of 7,690 M^{-1} cm^{-1} calculated from the extinction coefficients for tyrosine and tryptophan (Edelhoch, 1967). Zn and Cd contents were determined by flame atomic absorption using an Instrumentation Laboratories (Lexington, Massachusetts) IL157 spectrometer.

^{113}Cd NMR spectroscopy

^{113}Cd NMR spectroscopy was performed with a Bruker AM500 NMR spectrometer (110.93 MHz for ^{113}Cd) employing a 5-mm inverse probe. The UME6(111) concentration was 0.1 mM in 20 mM K-phosphate, 500 mM NaCl, at 35 °C and 5 °C. Spectra were processed using Felix 2.10 (Biosym Technologies, San Diego, California). A 30-Hz line broadening was applied during processing to reduce noise. The chemical shift reference is 0.7 M $^{113}Cd(ClO_4)_2$ at 25 °C.

Cloning of UME6(838)

The plasmid pPL5905 (Strich et al., 1994), containing the entire coding sequence of UME6, was digested with *Nhe* I and *Hin*d III (New England Biolabs, Beverly, Massachusetts). The resulting 2.8-kb fragment containing the sequence for the C-terminal 765 amino acids was gel-purified using Geneclean (Bio101, Vista, California). The plasmid pAR3039 (Rosenberg et al., 1987) was also digested with *Nhe* I and *Hin*d III, and the 4.1-kb fragment containing the phage T7 promoter was also gel purified. The purified fragments were ligated together using phage T4 ligase and electroporated into *E. coli* strain BL21(DE3) (Studier & Moffatt, 1986). Colonies were screened for appropriate restriction sites and sequenced with Sequenase 2.0 (U.S. Biochemical, Cleveland, Ohio). Correct constructs, pUME6(766), result in the expression of the protein sequence: Met-Ala^{72}... Lys^{836}. PCR primers were used to amplify the N-terminus of the UME6 gene from pPL5905. The upstream primer, ACCTT AACTCGCTAGCCTAGACAA, inserts an *Nhe* I site at the N-terminus of the gene, whereas the downstream primer, ACTTTCAAACTCGGCGAACCAGAC, anneals 93–116 bp downstream of the *Nhe* I site of UME6 without altering the gene sequence. The resulting 348-bp PCR product was digested by *Nhe* I to yield a 210-bp fragment that was ligated into the *Nhe* I site in pUME6(766). Plasmids with the correct sequence can be cut by *Bss*H II, and such plasmids were sequenced to confirm insertion of the fragment and determine its orientation. Correct constructs result in expression of a protein with the sequence: Met-Ala-Ser-Leu^{2}... Lys^{836} and are referred to as pUME6(838).

Overexpression and purification of UME6(838)

BL21(DE3) cells containing the plasmid pUME6(838) were grown at 37 °C in 1.5 L of LB medium to an OD_{600nm} of 0.5–0.7 and induced with 0.15 g/L IPTG. Immediately after induction, the growth temperature was switched to 20 °C and cells were grown for an additional 4 h. Cells were harvested and immediately resuspended in 25 mL of sonication buffer (20 mM Tris·HCl, pH 7.5, 300 mM NaCl, 10 mM BME, 1 mM $ZnCl_2$, 0.1% Tween-20, 0.17 mg/mL PMSF [Sigma]) and sonicated for 3 × 3 min, followed by centrifugation at 27,000 × *g* for 10 min. Polyethylenimine, 1.5 mL of 9%, pH 8.0 (Sigma), was added to the supernatant with stirring on ice to precipitate negatively charged proteins and nucleic acids. Additional PMSF was added, again to 0.17 mg/mL, and the mixture was centrifuged at 27,000 × *g* for 10 min. The supernatant was diluted fourfold with buffer B, which is the same as buffer A, but contains 0.1% Tween-20 (Sigma) and 0.17 mg/mL PMSF, and loaded onto a 60-mL Trisacryl-SP column. The column was washed with three volumes of buffer B and bound proteins were eluted with 100 mL of buffer B + 1 M NaCl. Eluted protein was diluted fourfold with column buffer and loaded onto an ssDNA-cellulose column, washed with three volumes of buffer B + 250 mM NaCl, and eluted with a 250 mM to 1 M NaCl gradient in buffer B. Fractions containing UME6(838) were identified by staining of SDS-PAGE gels of column fractions with Coomassie brilliant blue R-250 and used without further purification.

CD

CD was performed on an AVIV (Lakewood, New Jersey) 62DS CD spectrometer at 25 °C using a cell with a 0.2-cm pathlength. Data points were collected every 0.5 nm, with signal averaging for 2.0 s/point at a bandwidth of 1 nm. Each scan was repeated five times, averaged, and smoothed with a 3-point running average function to reduce noise.

Construction of pI2U1, a plasmid containing the URS1 sequence from the IME2 gene

An oligonucleotide with the sequence CGCTCTAGATCCTTT TCTCGCGTTGTCCAATAATTTATGTTACGGCGGCTATT TGAGC and a primer TCGAGCTCAAATAGCCGCGCTAAC complementary to the 3′ end were synthesized by the Yale Pathology Department DNA synthesis facility, New Haven, Connecticut. This sequence contains a URS1 site and the associated T4C box (Bowdish & Mitchell, 1993) from just upstream of the IME2 promoter. Both oligonucleotides were phosphorylated with T4 polynucleotide kinase (Sambrook et al., 1989), annealed, and incubated overnight at 16 °C in the presence of T4 DNA ligase and 2 mM ATP. The resulting dimer was Klenow filled, gel purified, and digested with 20 units each of *Xba* I and *Xho* I. This *Xba* I/*Xho* I fragment was ligated into the vector pBEND5 (Zweib & Adhya, 1994) digested with *Xba* I and *Sal* I. *Sal* I digestion was used to destroy recircularized plasmid and the mixture was electroporated into *E. coli* strain JM101. One colony containing the correct *Eco*R I/*Hin*d III fragment was sequenced with Sequenase 2.0 (U.S. Biochemicals) and shown to have the correct IME2 URS1 fragment insertion.

Preparation of DNA fragments for gel retardation

A duplex containing the SPO13 URS1 sequence was synthesized (Yale Pathology Department DNA Synthesis Facility) with the following sequence:

5′-AATTCCTTTTTGTCGGCGGCTATTTCTGCA-3′

3′-GGAAAAACAGCCGCCGATAAAG-5′

The shorter strand was ^{32}P-labeled at the 5′-end as described above, a twofold excess of the unlabeled complementary strand

was added, and the strands were annealed by heating to 95 °C and slowly cooling to room temperature. The DNA fragment containing the IME2 URS1 sequence was prepared by digesting 0.3 pmol of plasmid pI2U1 for 2 h with *Xho* I. The resulting fragments were either separated on a 1% agarose gel using the MERmaid gel purification kit (Bio101) or used as a mixture of fragments without purification. Fragments were ^{32}P-labeled using standard procedures (Sambrook et al., 1989) and purified using Sephadex G-50 spun columns.

Gel retardation of URS1-containing and nonspecific DNA by UME6(111) and UME6(838)

UME6 protein was diluted in 10 mM $NaPO_4$, pH 7.5, 10 μM $ZnCl_2$, and 1 μg/μL acetylated BSA. The salt concentration in each reaction tube was adjusted to a final concentration of 115 mM NaCl. Binding reactions were carried out in the same buffer used for dilutions but containing 20% glycerol and 1 nM ^{32}P-labeled URS1-containing DNA. The labeled nucleic acid was mixed with diluted UME6 and incubated at 20 °C for 5 min before loading onto an 8% polyacrylamide gel in TB buffer (90 mM Tris-borate, pH 9.0) and electrophoresing for 45 min at 300 V. Gels were dried and exposed to Kodak XAR film overnight at room temperature. Band intensities were quantitated either by re-exposing the gel to a Fuji Bas-III phosphorimage plate and reading the plate with a Fuji Bas1000 phosphorimager, or quantitated directly from the autoradiogram using an LKB Ultroscan XL laser densitometer.

DNase I footprinting

An aliquot of UME6(838) containing 1 pmol of protein was added to 20 fmol of ^{32}P-labeled URS1 DNA duplex in 10 μL of 10 mM $NaPO_4$, pH 7.5, 10 μM $ZnCl_2$, 1 μg/μL acetylated BSA, and 3 nM unlabeled calf thymus competitor DNA. The sample was incubated at 20 °C for 5 min, and 1 μL of a solution containing 100 mM $MgCl_2$ and 50 mM $CaCl_2$ and 1 μL of DNase I (0.005 U/μL) was added. After 1 min, 10 μL of stop solution (1% SDS, 200 mM NaCl, 20 mM EDTA, 83% glycerol), 2 μL of 10 M NaOAc, and 2 μL of 20 μg/μL "carrier" tRNA were added. DNA was precipitated by addition of 60 μL of 95% EtOH and redissolved in 20 μL of formamide loading buffer. Equal amounts of radioactivity were loaded in control and footprint lanes and electrophoresed at 65 W for 2–3 h on 20% polyacrylamide/7 M urea sequencing gels. The gel was dried and exposed to a Fuji BAS-III phosphorimaging plate for 1–4 h, and the band intensities were quantitated on a Fuji BAS1000 phosphorimager. Phosphorimager data were converted to density cross-section format using the Macintosh program NIH Image and displayed using SigmaPlot (Jandel Scientific).

Acknowledgments

This work was supported by NIH grants DK09070 and GM21919 to J.E.C. and NIH grants GM29182 and HD19252 to R.E.E. This work is in partial fulfillment of the Ph.D. requirements for S.F.A., who was supported by a predoctoral fellowship from the Howard Hughes Medical Institute. Plasmid pBEND5 was obtained from Christian Zweib.

References

Baleja JD, Marmorstein R, Harrison SC, Wagner G. 1992. Solution structure of the DNA binding domain of Cd_2GAL4 from *S. cerevisiae*. *Nature 356*:450–453.

Bowdish KS, Mitchell AP. 1993. Bipartite structure of an early meiotic upstream activation sequence from *Saccharomyces cerevisiae*. *Mol Cell Biol 13*:2172–2181.

Bram RJ, Kornberg RD. 1985. Specific protein binding to far upstream activating sequences in polymerase II promoters. *Proc Natl Acad Sci USA 82*:43–47.

Buckingham LE, Wang H-T, Elder RT, McCarroll RM, Slater MR, Esposito RE. 1990. Nucleotide sequence and promoter analysis of SPO13, a meiosis-specific gene of *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA 87*:9406–9410.

Carey M, Kakidani H, Leatherwood J, Mostashari F, Ptashne M. 1989. An amino terminal fragment of GAL4 binds as a dimer. *J Mol Biol 209*: 423–432.

Coleman JE. 1993. Cadmium-113 nuclear magnetic resonance applied to metalloproteins. *Methods Enzymol 227*:16–43.

Coleman JE, Armitage IM, Chlebowski JF, Otvos JD, Schoot Uiterkamp AJM. 1979. Multinuclear NMR approaches to the solution structure of alkaline phosphatase: ^{13}C, ^{19}F, ^{31}P, and ^{113}Cd NMR. In: Shulman RG, ed. *Biological applications of magnetic resonance*. New York: Academic Press. pp 345–395.

de Rijcke M, Seneca S, Punyammalee B, Glansdorff N, Crabeel M. 1992. Characterization of the DNA target site for the yeast ARGR regulatory complex, a sequence able to mediate repression or induction by arginine. *Mol Cell Biol 12*:68–81.

Dhawale SS, Lane AC. 1993. Compilation of sequence-specific DNA-binding proteins implicated in transcriptional control in fungi. *Nucleic Acids Res 21*:5537–5546.

Dower WJ, Miller JE, Ragsdale CW. 1988. High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res 16*:6127–6145.

Edelhoch H. 1967. Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry 6*:1948–1954.

Gadhavi PL, Davis AL, Povey JF, Keeler J, Laue ED. 1991. Polypeptide-metal cluster connectivities in Cd(II)GAL4. *FEBS Lett 281*:223–226.

Gardner KH, Coleman JE. 1994. ^{113}Cd-^{1}H heteroTOCSY: A method for determining metal-protein connectivities. *J Biomol NMR 4*:761–774.

Gardner KH, Anderson SF, Coleman JE. 1995. Solution structure of the *Kluveromyces lactis* LAC9 Cd_2Cys_6 DNA-binding domain. *Nature Structural Biol 2*:October 1995.

Gardner KH, Pan T, Narula S, Rivera E, Coleman JE. 1991. Structure of the binuclear metal-binding site in the GAL4 transcription factor. *Biochemistry 30*:11291–11302.

Giniger E, Varnum SM, Ptashne M. 1985. Specific DNA binding of GAL4, a positive regulatory protein of yeast. *Cell 40*:767–774.

Greenfield N, Fasman GD. 1969. Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry 8*:4108–4116.

Johnston M. 1987. A model fungal gene regulatory mechanism: The GAL genes of *Saccharomyces cerevisiae*. *Microbiol Rev 51*:458–476.

Kraulis PJ, Raine ARC, Gadhavi PL, Laue ED. 1992. Structure of the DNA-binding domain of zinc GAL4. *Nature 356*:448–450.

Kyte J, Doolittle RF. 1982. A simple method for displaying the hydropathic character of a protein. *J Mol Biol 157*:105–132.

Marmorstein R, Carey M, Ptashne M, Harrison SC. 1992. DNA recognition by GAL4: Structure of a protein-DNA complex. *Nature 356*:408–414.

Marmorstein R, Harrison SC. 1994. Crystal structure of a PPR1-DNA complex: DNA recognition by proteins containing a Zn_2Cys_6 binuclear cluster. *Genes & Dev 8*:2504–2512.

Mitchell AP. 1994. Control of meiotic gene expression in *Saccharomyces cerevisiae*. *Microbiol Rev 58*:56–70.

Mitchell AP, Driscoll SE, Smith HE. 1990. Positive control of sporulation-specific genes by the IME1 and IME2 products in *Saccharomyces cerevisiae*. *Mol Cell Biol 10*:2104–2110.

Pan T, Coleman JE. 1989. Structure and function of the Zn(II) binding site within the DNA-binding domain of the GAL4 transcription factor. *Proc Natl Acad Sci USA 86*:3145–3149.

Pan T, Coleman JE. 1990a. The DNA binding domain of GAL4 forms a binuclear metal ion complex. *Biochemistry 29*:3023–3029.

Pan T, Coleman JE. 1990b. GAL4 transcription factor is not a "zinc finger" but forms a $Zn(II)_2Cys_6$ binuclear cluster. *Proc Natl Acad Sci USA 87*:2077–2081.

Pan T, Halvorsen YD, Dickson RC, Coleman JE. 1990. The transcription factor LAC9 from *Kluyveromyces lactis* like GAL4 from *Saccharomy-*

ces cerevisiae forms a $Zn(II)_2Cys_6$ binuclear cluster. *J Biol Chem 265*:21427–21429.

Park HD, Luche RM, Cooper TG. 1992. The yeast UME6 gene product is required for transcriptional repression mediated by the CAR1 URS1 repressor binding site. *Nucleic Acids Res 20*:1909–1915.

Rodgers KK, Coleman JE. 1994. DNA binding and bending by the transcription factors GAL4(62*) and GAL4(149*). *Protein Sci 3*:608–619.

Rosenberg AH, Lade BN, Chui DS, Lin SW, Dunn JJ, Studier WF. 1987. Vectors for selective expression of cloned DNA's by T7 RNA polymerase. *Gene 56*:125–135.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning–A laboratory manual 2nd ed.* Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Saxena VP, Wetlaufer DB. 1971. A new basis for interpreting the circular dichroic spectra of proteins. *Proc Natl Acad Sci USA 68*:969–972.

Shirakawa M, Fairbrother WJ, Serikawa Y, Ohkubo T, Kyogoku Y, Wright PE. 1993. Assignment of ^{1}H, ^{15}N, and ^{13}C resonances, identification of elements of secondary structure and determination of the global fold of the DNA-binding domain of GAL4. *Biochemistry 32*:2144–2153.

Smith HE, Mitchell AP. 1989. A transcriptional cascade governs entry into meiosis in *Saccharomyces cerevisiae*. *Mol Cell Biol 9*:2142–2152.

Strich R, Slater MR, Esposito RE. 1989. Identification of negative regulatory genes that govern the expression of early meiotic genes in yeast. *Proc Natl Acad Sci USA 86*:10018–10022.

Strich R, Surosky RT, Steber C, Dubois E, Messenguy F, Esposito RE. 1994. UME6 is a key regulator of nitrogen repression and meiotic development. *Genes & Dev 8*:796–810.

Studier FW, Moffatt BA. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol 189*:113–130.

Sumrada RA, Cooper TG. 1987. Ubiquitous upstream repression sequences control activation of the inducible arginase gene in yeast. *Proc Natl Acad Sci USA 84*:3997–4001.

Wishart DS, Boyko RF, Willard L, Richards FM, Sykes BD. 1994. SEQSEE: A comprehensive program suite for protein sequence analysis. *Comput Appl Biosci 10*:121–132.

Zweib C, Adhya S. 1994. Improved plasmid vectors for the analysis of protein-induced DNA bending. *Methods Mol Biol 30*:281–294.