Flexibility of the yeast $\alpha 2$ repressor enables it to occupy the ends of its operator, leaving the center free

Robert T. Sauer,¹ Dana L. Smith, and Alexander D. Johnson

Departments of Microbiology and Biochemistry/Biophysics, University of California, San Francisco, California 94143 USA

The yeast $\alpha 2$ protein, the product of the *MAT $\alpha 2$* gene, is a regulator of yeast cell type; it turns off transcription of the *a*-specific genes by binding to an operator located upstream of each gene. In this paper we describe the domain structure, subunit organization, and some unusual features of the way this protein contacts its operator. We show that the protein is folded into two domains. The carboxy-terminal domain binds specifically to the operator; the amino-terminal domain contains dimerization contacts. The $\alpha 2$ dimer differs from those of the phage repressors in that it is flexible and therefore is able to bind tightly to differently spaced operator half-sites. In the natural operator, the centers of the operator half-sites are two and one-half turns of DNA apart, exposing them on opposite sides of the DNA helix. We show that the design of $\alpha 2$ allows a dimer to reach across its operator such that it occupies the two half-sites but leaves the middle of the operator available to other proteins.

[Key Words: DNA-protein interaction; repressor; gene expression; homeo domain]

Received March 29, 1988; revised version accepted May 13, 1988.

The yeast $\alpha 2$ protein is related by sequence (and presumably structure) and by function (the determination of cell type) to a large group of proteins that contain the so-called homeo domain (for review, see Gehring 1987). These proteins are all thought to act, at least in part, by binding to specific DNA sequences and modulating the expression of nearby target genes (see e.g., Laughon and Scott 1984; Desplan et al. 1985). For the case of $\alpha 2$, we know that this protein recognizes an operator located upstream of each of its target genes (the *a*-specific genes) and that it turns off transcription of these genes by binding to these operators (see Fig. 1; Johnson and Herskowitz 1985).

In this paper we utilize $\alpha 2$ purified to homogeneity from an overproducing strain of *Escherichia coli* to determine the domain structure, subunit arrangement, and functional organization of $\alpha 2$. We show that the design of $\alpha 2$ and its operator allows access of proteins to the center of the operator even when $\alpha 2$ is bound.

Results

Overexpression and purification of the $\alpha 2$ protein

As described in Methods, the $\alpha 2$ protein was overexpressed in *E. coli* and purified to apparent homogeneity. An SDS gel profile of the purified protein is shown in Figure 2, lane 1. As predicted from the DNA sequence (Astell et al. 1981), $\alpha 2$ has a monomer molecular weight

of 24,000, its amino-terminal sequence is Met-Asn-Lys-Ile..., and the purified protein exhibits the expected amino acid composition (see Methods). The presence of the amino-terminal Met is consistent with the fact that the *E. coli* methionine aminopeptidase fails to cleave before an asparagine (see Flinta et al. 1986; Ben-Bassat and Bauer 1987). We do not yet know whether the amino-terminal methionine is present in $\alpha 2$ produced from yeast, but the protein produced in *E. coli* comigrates with $\alpha 2$ from yeast on SDS gels (A. Vershon, unpubl.). This suggests that no drastic modifications of $\alpha 2$ occur in yeast.

Domain structure of the $\alpha 2$ protein

Cleavage of the purified $\alpha 2$ protein with chymotrypsin generates two major fragments. As shown in Figure 2B, the yield of these fragments remains relatively constant between 10 and 640 min of digestion. This resistance to further proteolysis indicates that these fragments are folded stably and correspond to structural domains of $\alpha 2$. The chymotryptic fragments of $\alpha 2$ were purified, and their amino-terminal sequences and amino acid compositions were determined (see Methods). This analysis showed that the larger fragment consists of amino acids 1–102 and the smaller fragment consists of amino acids 132–210 of $\alpha 2$ (210 residues total). We will refer to these fragments as the amino-terminal and carboxy-terminal fragments, respectively. (Close inspection of Fig. 2 reveals a minor digestion product that migrates just ahead of the carboxy-terminal fragment; this minor product appears at early times of digestion but disappears at late

¹On sabbatical leave from the Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 USA

Figure 1. Control of the α -specific genes by $\alpha 2$. (Top) $\alpha 2$ turns off transcription of the five known α -specific genes by binding to an operator located upstream of each gene; for the cases of the *STE6*, *BAR1*, and *STE2* genes, the binding has been demonstrated biochemically (Johnson and Herskowitz 1985; L. Evnin and A. Johnson, unpubl.). Interaction of $\alpha 2$ with the *MFa1* and *MFa2* sequences has been inferred on the basis of strong sequence conservation. The numbers beneath each operator refer to the approximate distance from the center of the operator to the start of the structural gene. (Bottom) The sequences of the five operators. The sixth line shows a rough consensus sequence. The seventh and eighth lines indicate that the operator contains two different axes of approximate twofold symmetry; the two axes are offset by 1/2 bp. Taken from Johnson and Herskowitz (1985), this figure includes sequence information from the following sources: *STE6* (Wilson and Herskowitz 1986) *BAR1* (Kronstad et al. 1987), *STE2* (Burkholder and Hartwell 1985; Miller et al. 1985), *MFa1* and *MFa2* (Brake et al. 1985).

times. We also purified this fragment and determined its amino-terminal sequence and composition; it consists of amino acids 132–205. Because of low yields, this fragment was not studied further.)

Subunit structure of the intact $\alpha 2$ protein

Purified $\alpha 2$ exists as a covalently linked dimer, with the linkage occurring through disulfide bonds between amino-terminal domains. This conclusion is based on the observation that the apparent size of $\alpha 2$ on SDS gels doubles if the protein is not treated with reducing agents prior to electrophoresis. Lane 1 of Figure 2A shows that when $\alpha 2$ was denatured by heating in SDS and 2-mercaptoethanol, it migrated with the expected molecular weight of 24,000. Lane 2 of Figure 2A shows that when the 2-mercaptoethanol was omitted from the sample buffer, the purified protein migrated with an apparent molecular weight of approximately 45,000.

Next, we used gel filtration chromatography and sedimentation velocity centrifugation experiments to examine the apparent molecular weight of reduced and disulfide-linked $\alpha 2$ under native conditions. These results are shown in Figures 3 and 4. Reduced $\alpha 2$ migrates through the sizing column with an apparent molecular weight of 35,000 and sediments through a sucrose gradient with an apparent molecular weight of 17,000. These values can be compared to the expected molecular weight of an $\alpha 2$ monomer, 24,000. The simplest interpretation of these data is that reduced $\alpha 2$ is a monomer with a somewhat elongated shape. A molecule of this type would migrate faster than expected through a sizing column and slower than expected through a sucrose gradient. Further evidence that reduced $\alpha 2$ is, in fact, mon-

omeric comes from sedimentation equilibrium experiments performed by Bill Werner and Howard Schachman at University of California, Berkeley. They observed that, at equilibrium, the reduced $\alpha 2$ was dis-

Figure 2. SDS-gel profiles of intact $\alpha 2$ and chymotryptic fragments of $\alpha 2$. (A) Disulfide-linked and reduced $\alpha 2$. (Lane 1) Purified $\alpha 2$ electrophoresed through a 12% acrylamide gel, according to Laemmli (1970). 2-Mercaptoethanol was included in the sample buffer, and the sample was heated at 90°C for 2 min prior to electrophoresis. (Lane 2) An aliquot from the same batch of $\alpha 2$ treated in the same way, except that the 2-mercaptoethanol was omitted from the sample buffer. (B) Generation of chymotryptic fragments of $\alpha 2$. (Lane 1) Purified $\alpha 2$; (lane 2) blank; (lane 3–11) purified $\alpha 2$ after treatment with chymotrypsin (E : S = 1 : 6000) for 2.5, 5, 10, 20, 40, 80, 160, 320, and 640 min, respectively. All samples were heated in 2-mercaptoethanol prior to electrophoresis through a 12% SDS gel.

Figure 3. Mobilities of reduced and disulfide-linked $\alpha 2$ on a superose-12 gel filtration column. The molecular weights of the standards are IgG (158,000), BSA (68,000), ovalbumin (43,000), myoglobin (17,000) and cytochrome c (12,000). The relative mobility of IgG was set at 1.0, and that of myoglobin was set at 0. For this experiment, the chromatography was carried out in T buffer plus 0.5 M NaCl.

tributed along the cell with a profile corresponding to a single species of 25 kD (not shown).

The disulfide-linked form of $\alpha 2$ protein migrates through the sizing column with an apparent molecular weight of 90,000 (Fig. 3) and sediments through a sucrose gradient with an apparent molecular weight of 25,000 (Fig. 4). These data are most easily explained by the idea that the disulfide-linked form of $\alpha 2$ is an elongated dimer. Thus, the simplest explanation of the SDS gel, gel filtration, and sedimentation velocity experiments is as follows: The major solution form of $\alpha 2$ is the disulfide-linked dimer, which is highly elongated. Dimers of $\alpha 2$ are broken down into monomers by reducing the disulfide bonds. In the remaining discussion, we use the term 'dimer' to refer to the disulfide-linked species of $\alpha 2$ and 'monomer' to refer to the reduced species.

Two lines of argument indicate that the disulfide bonds form between the amino-terminal domains of $\alpha 2$. First, the only cysteines in the molecule are located in the amino-terminal domain, at positions 33 and 34. Second, after the chymotryptic digestion, which was carried out on $\alpha 2$ dimers, the amino-terminal domain was recovered as a disulfide-linked dimer, whereas the carboxy-terminal domain was recovered as a monomer (not shown). We note that in principle, one or two disulfide bonds could form between amino-terminal domains; we have not determined this number, nor have we determined which of the two cysteines is actually involved.

Operator binding by $\alpha 2$ dimers, monomers, and chymotryptic fragments

We now turn to experiments that examine the interaction of purified $\alpha 2$ with its operator. Figure 5A shows a

DNase I protection experiment where increasing concentrations of $\alpha 2$ dimers have been added to an end-labeled DNA fragment that bears a synthetic $\alpha 2$ operator. This synthetic operator, which functions *in vivo* (Johnson and Herskowitz 1985), has the sequence of the operator found upstream of the *STE6* gene (Wilson and Herskowitz 1986; for the sequence, see the top line of Fig. 7). Three conclusions can be drawn from the protection experiment. First, the purified $\alpha 2$ protein binds specifically to the operator. This result was expected because a full-length $\alpha 2$ - β -galactosidase fusion protein, purified from yeast, had previously been shown to bind to this sequence (Johnson and Herskowitz 1985). Second, the purified $\alpha 2$ protein protects the two end portions of the operator from DNase I attack, leaving the middle of the operator accessible to DNase I. These two protected sequences are related by an approximate twofold axis of symmetry (see Figs. 1 and 6, below); we refer to each sequence as a half-site. Third, the two half-sites are occupied at about the same concentration of $\alpha 2$ dimers, ~ 3 nM.

To map the portion of the operator contacted by $\alpha 2$ more accurately, we carried out a series of hydroxyl radical protection experiments, using the procedure described by Tullius and Dombroski (1986). The hydroxyl radical probe is about the size of a water molecule, and it attacks ribose positions along the DNA backbone. The results of one such protection experiment are shown in Figure 5B; three short areas of protection are seen. The results of this experiment and a similar one performed to probe the complementary DNA strand are summarized in Figure 6. The top of Figure 6 shows the backbone positions protected from DNase I (indicated by brackets) and hydroxyl radical attack (indicated by solid circles) by the purified $\alpha 2$ protein. The bottom of Figure 6 shows the hydroxyl radical results displayed on a projection of

Figure 4. Relative sedimentation of reduced and disulfide-linked $\alpha 2$. Molecular weights of the standards and the explanation of the x-axis values are given in the legend of Fig. 3. This experiment was carried out in T buffer plus 0.2 M NaCl.

the operator sequence arranged as B-form DNA. The projection diagram shows that $\alpha 2$ contacts two patches of DNA, each consisting of a portion of a major groove and an adjoining portion of the minor groove. The centers of the two patches are approximately two and one-half turns of DNA apart and therefore reside on opposite faces of the DNA helix, as indicated in the diagram. We also performed a series of dimethylsulfate (DMS) protection experiments and found that $\alpha 2$ protected the N-7 positions of two symmetrically located Gs and failed to protect the N-3 positions of any As (Figs. 5C and 6). This result suggests that the contacts made by this protein reside primarily in the major groove of each half-site. The protected Gs are indicated in Figure 6 as open circles.

The results summarized above indicate that $\alpha 2$ does not contact the middle of the operator. To test this idea

experimentally, we compared the affinity of $\alpha 2$ dimers for the wild-type *STE6* operator with that for an operator site that contained a 7-bp substitution in its center (for the sequence, see the second line of Fig. 7). This variant operator therefore retains the two half-sites with their normal spacing. We found that $\alpha 2$ had the same affinity for the mutant as the wild-type operator site (not shown), confirming that the sequence of the middle of the operator is not directly involved in $\alpha 2$ recognition.

Figure 5. Interaction of purified $\alpha 2$ dimers with the *STE6* operator. (A) Protection of the *STE6* operator from DNase I attack by $\alpha 2$ dimers. The DNA is an 80-bp fragment that contains the synthetic *STE6* operator (top sequence of Fig. 7) cloned into the *Sall* site of the pUC18 polylinker (Viera and Messing 1982). For this experiment, the top strand is labeled at the right end, according to the orientation of the sequence in Figs. 6 and 7. The sequences protected from DNase I attack are shown in Fig. 6. The concentration of $\alpha 2$ dimers in each lane is as follows: (Lane 1) No protein; (lane 2) 0.1 nM; (lane 3) 0.3 nM; (lane 4) 1 nM; (lane 5) 3 nM; (lane 6) 10 nM; (lane 7) 30 nM. (B) Protection of the *STE6* operator from hydroxyl radical attack by $\alpha 2$ dimers. The DNA fragment is the same as that in A. (Lane 1) 30 nM $\alpha 2$ dimers; (lane 2) no protein. The results are summarized in Fig. 6. (C) Protection of the *STE6* operator from DMS attack by $\alpha 2$ dimers. The DNA fragment is the same as that in A. (Lane 1) No protein; (lane 2) 30 nM $\alpha 2$ dimers. The results are summarized in Fig. 6.

Figure 6. Summary of DNase I, hydroxyl radical, and DMS protection results. (Top) The *STE6* operator cloned into the polylinker of pUC18. The brackets indicate the regions of DNA protected from DNase I attack by purified $\alpha 2$ protein. (●) Backbone positions protected from hydroxyl radical attack by $\alpha 2$; (○) the two Gs protected from attack by DMS. The arrows beneath the sequence indicate the approximate twofold symmetry shown by the sequences of the protected regions (see also Fig. 1). (Bottom) The hydroxyl radical and DMS protection results displayed on a split projection of the operator as B-form DNA. Note that the pattern of protected positions is roughly symmetric around the same axis of symmetry as the base sequence is.

To test whether the natural spacing of the two half-sites (two and one-half turns apart) is a requisite for $\alpha 2$ binding, we synthesized a variant of the *STE6* operator that had the central 13 bp deleted (for the sequence, see the fourth line of Fig. 7). In this variant operator, the two half-sites are directly adjacent and lie on the same face of the DNA helix. We found that $\alpha 2$ dimers did indeed bind to this variant operator, occupying both half-sites at the same $\alpha 2$ concentration (Fig. 8). Surprisingly, $\alpha 2$ dimers bound to the variant with an affinity similar to, or slightly higher than, the affinity for the wild-type operator.

We next performed an experiment to determine whether two half-sites bind $\alpha 2$ dimers more strongly than a single half-site does. As shown by the DNase I protection experiment of Figure 9, $\alpha 2$ dimers did bind specifically to a single half-site (for the sequence, see the third line of Fig. 7), albeit with an affinity ~ 10 -fold lower than that for the wild-type operator (line 1 of Table 1).

The results so far suggest a simple model for $\alpha 2$ dimers binding to an operator (see Fig. 11). Each monomer contacts a half-site, and the dimer is sufficiently flexible to contact both half-sites regardless of

the spacing. A prediction of this model is that $\alpha 2$ monomers (formed by reducing the disulfide bonds) should bind to the operator less tightly than $\alpha 2$ dimers do. This prediction was tested and found to be true; $\alpha 2$ monomers specifically recognized the operator, but with an affinity ~ 20 -fold lower than that of $\alpha 2$ dimers (line 2 of Table 1).

Finally, we determined whether either of the two isolated domains of $\alpha 2$ retained the ability to recognize the operator. We anticipated that the carboxy-terminal domain would bind to the operator, because previous work utilizing deletion derivatives of an $\alpha 2$ - β -galactosidase fusion protein (Hall and Johnson 1987) showed that the DNA-binding portion of $\alpha 2$ lay at the carboxyl-terminus. As expected, DNase I protection experiments showed that the purified amino-terminal domain failed to bind operator DNA, whereas the purified carboxy-terminal domain showed protection of the operator indistinguishable from that of the intact $\alpha 2$ (Fig. 10). The affinity of the isolated carboxy-terminal domain, however, was about 20-fold lower than that of the intact $\alpha 2$ dimers. The isolated carboxy-terminal domain also bound to the half-operator; in contrast to the intact $\alpha 2$ dimers, the affinity of the carboxy-terminal fragment for the wild-type

Figure 7. DNA sequences of the *STE6* operator and three variants. The oligonucleotides were synthesized, cloned into the *SalI* site of pUC18, isolated as longer DNA fragments, and used for the protection experiments of Figures 5, 8, 9, and 10 and Table 1. To the right of each operator sequence is a schematic description of it; the hatched sections indicate the half-sites recognized by $\alpha 2$.

Figure 8. Binding of $\alpha 2$ dimers to the operator with the 13-bp central deletion. Binding of $\alpha 2$ to the variant operator (last sequence of Fig. 7) monitored by DNase I protection. Conditions and concentration of $\alpha 2$ are the same as those of Figure 5A; the experiment of lane 8 contains $\alpha 2$ at 100 nM.

and half-operator variant was approximately the same. These results are listed in line 3 of Table 1.

Discussion

Domain and subunit structure of $\alpha 2$

Our results show that the $\alpha 2$ protein consists of two structural domains. The isolated carboxy-terminal domain binds specifically to the operator; the isolated amino-terminal domain forms dimers where the monomers are linked by disulfide bonds. In standard buffers, the predominant solution form of intact $\alpha 2$ is the disulfide-bonded dimer. The migration rates of $\alpha 2$ dimers through gel filtration columns and sucrose gradients indicate that $\alpha 2$ is elongated.

Interaction of $\alpha 2$ with the operator

$\alpha 2$ dimers bind moderately tightly to the operator ($K_d = 3$ nM) and contact two patches of DNA, one located at each end of the operator. The centers of these two patches are separated by 25 bp, thus positioning them on opposite faces of the DNA helix (see Fig. 6). As

a result, the middle of the operator is accessible to DNase I and hydroxyl radical probes even when $\alpha 2$ is bound. The affinities with which the $\alpha 2$ dimer, monomer, and the carboxy-terminal domain bind to the wild-type operator and to a half-operator (Table 1), indicate that a single $\alpha 2$ dimer contacts both portions of the operator, as shown in the model of Figure 11. Thus, the binding energies to the two half-sites of the operator are coupled, and the coupling is mediated through the disulfide-linked interactions between amino-terminal domains. We note that several alternative models are also consistent with our data. For example, two dimers could bind to the operator in a highly cooperative manner, with the cooperativity dependent upon the disulfide bonds being present. However, the model presented in Figure 10 is by far the most reasonable and economical summary of the results.

The fact that $\alpha 2$ dimers bind the natural operator (where the center-to-center spacing of the half-sites is two and one-half turns) with the same affinity as they bind the 13-bp deleted operator (where the center-to-center spacing is about 1 turn) indicates that the $\alpha 2$ dimer has considerable flexibility. This flexibility appears to allow the dimer to recognize the two half-sites

Figure 9. Binding of $\alpha 2$ dimers to the half-operator variant, shown in the third line of Figure 7, was monitored by DNase I protection. Conditions are the same as those of Fig. 5A. Concentrations of $\alpha 2$ dimers are as follows: (Lane 1) No protein; (lane 2) 0.3 nM; (lane 3) 1 nM; (lane 4) 3 nM; (lane 5) 10 nM; (lane 6) 30 nM; (lane 7) 100 nM; (lane 8) 300 nM.

Table 1. Relative dissociation constants of $\alpha 2$ dimers, monomers, and the isolated carboxy-terminal domain for the wild-type *STE6* operator and the half-operator variant

	<i>STE6</i> operator	Half-operator
$\alpha 2$ dimers	3	20
$\alpha 2$ monomers	60	200
Carboxy-terminal domain	120	120

The entries give the nanomolar concentrations of different species of $\alpha 2$ required to half-saturate the *STE6* operator (column 1) or the half-operator variant (column 2). The values were obtained by carrying out a 10-point titration curve, using DNase I protection to monitor the occupancy (see Figs. 5, 8, 9, and 10). The concentration of protein required to occupy one-half of the DNA molecules was then determined by inspection and densitometry of the autoradiogram. The values are averages of at least three individual determinations and are believed to be accurate to within a factor of 2. The values are given in moles/liter of the appropriate species; therefore, the numbers in the first row are expressed in moles/liter dimers, those in the second row are moles/liter monomer, and those in the third are moles/liter monomers of the carboxy-terminal fragment.

when they are exposed on the same side of the DNA or on opposite sides of the DNA. We note that the 13-bp deleted operator bears the same half-site spacing as does the $\alpha 1/\alpha 2$ operator, a sequence bound by the combination of $\alpha 1$ and $\alpha 2$ (Goutte and Johnson 1988). Thus, this flexibility of $\alpha 2$ may be important for its ability to recognize both types of operators.

Dimerization

The results and models presented above raise a number of questions, several of which we will consider. First, are the disulfides that hold $\alpha 2$ dimers together in vitro relevant in vivo? This question arises, in part, because the intracellular environment of most cells strongly favors the reduced state. As a consequence, most known inter- and intrachain disulfide linkages are found on secreted proteins, although several examples of intracellular proteins that contain disulfide bonds are known. As for the $\alpha 2$ disulfide bonds, several possibilities exist. First, they may actually form in vivo. A preliminary set of experiments (A. Vershon, unpubl.) indicates that the major form of $\alpha 2$ in yeast is not disulfide linked, but it is still possible that the disulfides form only when two monomers of $\alpha 2$ are bound to the operator and are therefore in close proximity. A second possibility is that the cysteines serve in vivo as binding sites for a metal and that this metal holds the subunits of $\alpha 2$ together. As mentioned above, an $\alpha 2$ monomer contains two cysteines, at positions 33 and 34. Thus, two monomers of $\alpha 2$ could, in principle, be held together by a metal bound to two pairs of cysteines. This type of arrangement has been proposed recently to mediate dimerization of the HIV *tat* protein (Frankel et al. 1988). We also note that in metallothionein, several cases are known where two adjacent cysteines are found bonded to a metal (for review, see Hamer 1986). A third possibility is that the cysteines play no role at all in vivo, and the formation of the disul-

Figure 10. Interaction of the isolated carboxy-terminal domain of $\alpha 2$ with the *STE6* operator. Conditions are the same as for the experiment of Fig. 5A, except the purified carboxy-terminal fragment was used instead of the intact $\alpha 2$. (Lane 1) No protein; (lane 2) 0.2 nM; (lane 3) 0.6 nM; (lane 4) 2 nM; (lane 5) 6 nM; (lane 6) 20 nM; (lane 7) 60 nM; (lane 8) 200 nM; (lane 9) 600 nM. The concentrations are expressed as moles/liter of monomer.

fides is entirely artifactual. For example the amino-terminal domains could interact weakly, and formation of the disulfide could trap molecules in the dimeric state in vitro. Regardless of the role of the cysteines, we feel that the twofold symmetry of the operator, combined with the fact that $\alpha 2$ dimers bind more tightly to the operator than monomers do, suggests that the DNA-bound form of $\alpha 2$ in vivo is the dimer.

Figure 11. Summary of $\alpha 2$ domain and subunit organization. Our working model shows how $\alpha 2$ dimers interact with the operator. The two patches of the operator contacted by $\alpha 2$ were identified by the hydroxyl radical experiment of Fig. 5. The rationale for depicting a single dimer interacting with both half-sites is described in the text.

Operator structure

A second issue raised by our results concerns the function of the middle of the $\alpha 2$ operator. As shown above, $\alpha 2$ does not contact this part of the operator, nor does $\alpha 2$ require the middle sequence to bind. However, the middle of the operator is highly conserved among the five known operators of *a*-specific genes (Fig. 1). Keleher et al. (1988) have recently identified a protein, called GRM, which binds to the middle of the $\alpha 2$ operator. The operator binding of GRM and $\alpha 2$ is cooperative, and the combined action of both proteins is required to bring about repression in vivo. Apparently, the flexible and elongated structure of $\alpha 2$ enables it to straddle GRM yet still allow both proteins to bind to the operator.

Functions of the domains of $\alpha 2$

We now turn to a more detailed discussion of the functions of the individual domains of $\alpha 2$. As shown above, the isolated carboxy-terminal domain retains the ability to recognize the operator. This domain (amino acids 132–210) contains extensive homology to the *Drosophila* homeo domain sequences (Laughon and Scott 1984; Shepherd et al. 1984). According to a sequence alignment proposed by Pat O'Farrell (unpubl.), positions 136–190 of $\alpha 2$ show significant homology to the *Drosophila* homeo domain sequences. Our results strongly suggest that the homeo domain sequence of the $\alpha 2$ protein folds into a compact structure that binds specifically to DNA. Two other studies of the $\alpha 2$ protein are consistent with this view. Using deletion derivatives of the $\alpha 2$ protein, Hall and Johnson (1987) showed, that the operator binding region of the molecule lay between positions 130 and 204. Porter and Smith (1986) showed that a number of amino acid substitutions between positions 175 and 191 of $\alpha 2$ led to a loss of $\alpha 2$ function, as measured in vivo.

What are the functions of the amino-terminal domain of $\alpha 2$? One function, revealed by this study, is to provide a set of $\alpha 2$ dimerization contacts. A second function of this domain may be to contact the cell-type regulatory protein, *a1*, the product of the *MATa1* gene (Hall and Johnson 1987; Goutte and Johnson 1988). The *a1* protein has no intrinsic activity, yet in combination, *a1* and $\alpha 2$ act to repress a set of genes that is distinct from the set of *a*-specific genes, which $\alpha 2$ represses in the absence of *a1* (for reviews, see Herskowitz and Oshima 1981; Sprague et al. 1983; Nasmyth and Shore 1987). A third function of the amino-terminal domain of $\alpha 2$ is to contact the GRM protein; as discussed above, this protein binds to the center of the operator and serves as a corepressor with $\alpha 2$ (Keleher et al., 1988). A fourth function of the amino-terminal domain, identified by Hall et al. (1984), may be to provide a nuclear transport signal.

These ideas about the structural organization of $\alpha 2$ are similar in principle to those advanced to rationalize the domain structure of a number of prokaryotic repressors and activators. For these cases, the DNA-binding functions are often sequestered in one domain, whereas the functions that modulate regulatory activity (binding of

inducers or corepressors, proteolytic cleavage, interactions with other proteins) reside in a separate domain.

Methods

Buffers

Lysis buffer (LB) contains 100 mM Tris-HCl (pH 8), 1 mM EDTA, 10 mM 2-mercaptoethanol, 500 mM NaCl, 2 M urea, and 0.1% (vol/vol) NP-40. U buffer contains 50 mM Tris-HCl (pH 8), 1 mM EDTA, 10 mM 2-mercaptoethanol, and 5 M urea. S buffer contains 50 mM Tris-HCl (pH 8), 1 mM EDTA, and 0.28 mM EDTA. T-buffer contains 50 mM Tris-HCl (pH 8) and 1 mM EDTA. Buffers containing urea were prepared immediately prior to use.

Construction of a pTAC- $\alpha 2$ overproducing strain

Plasmid pDS1 is a pTAC12 derivative that bears the *MATa2* gene from the yeast *Saccharomyces cerevisiae* under transcriptional control of the strong *E. coli* promoter, *P^{lac}*. To construct this plasmid, an 800 bp *Bgl*III–*Hpa*I fragment extending from the seventh codon of $\alpha 2$ (see Hall and Johnson 1987) through the remainder of the structural gene was purified and mixed with *Pvu*II-cut plasmid pTAC12 (Amann et al. 1983) and a synthetic oligonucleotide,

```
MetAsnLysIleProIleLys...
5'-CGTATGAATAAAATACCCATTAAA
3'-GCATACTTATTTTATGGGTAATTTCTAG
```

which provided the amino-terminal codons of $\alpha 2$. Following ligation and transformation of *E. coli* strain SY903.1 [*recA1*, *srl*::Tn10, Δ *lacpro*, *argEam*, *rif*, *nal*, *araD*/*F'**lacQ1*, *lacZ*::Tn5], a candidate plasmid with the expected structure was identified by hybridization and DNA sequencing. Following induction with isopropyl- β -D-thiogalactoside (IPTG), strains containing this plasmid produce $\alpha 2$ at a level of 1–2% of their total cell protein.

Protein purification

Strain SY903.1/pDS1 was grown at 30°C in LB to an absorbance at 600 nm of 1. IPTG was then added to a concentration of 0.5 mM and growth was continued for 2 hr. The cells were harvested by centrifugation and stored frozen at –70°C.

All steps of purification were performed at 2–4 °C and were monitored by SDS–gel electrophoresis. Cells (26.6 g) were thawed, resuspended in 266 ml of lysis buffer, and sonicated until the *A*₆₀₀ was reduced fivefold. Polyethyleneimine [16 ml of a 10% (vol/vol) solution] was added; the mixture was stirred for 10 min and centrifuged (8000 rpm; 15 min) in a Sorval GSA rotor. The pellet was discarded. Ammonium sulfate (180 ml of a saturated solution in T buffer) was added to the supernatant (277 ml), the mixture was left for 15 min, and the pellet fraction was collected by centrifugation (conditions as described above). The 0–40% ammonium sulfate pellet was resuspended in 70 ml of U buffer, desalted into fresh U buffer on a 500-ml column of BioRad P6D, and passed through a 30-ml column of DEAE–Sephacel equilibrated in U buffer. The flow through fraction was loaded onto a 20-ml column of SP–Sephadex (C50), and a gradient (200 ml total) was run from U buffer to U buffer plus 500 mM NaCl. Fractions containing $\alpha 2$ were pooled (35 ml total) and dialyzed against several changes of S buffer plus 500 mM NaCl. The SP pool was divided into three portions, which were chromatographed on a 2.5 × 82-cm column of Sephacryl S-200 equilibrated in S buffer plus 500 mM NaCl.

Generation and purification of chymotryptic fragments

Approximately 5 mg of purified $\alpha 2$ protein (0.3 mg/ml in T buffer plus 50 mM NaCl) was digested with α -chymotrypsin [E : S = 1 : 600 (wt/wt)] for 3 hr at 21°C. The reaction was stopped by the addition of 0.1 volume of a fresh solution of phenylmethylsulfonyl fluoride (PMSF) (50 mM in EtOH), and fragments were purified by a combination of ion-exchange chromatography on SP-Sephadex and gel filtration on Sephadex G-75.

Protein chemistry

Samples for amino acid analysis were hydrolyzed in 6 N HCl for 24 hr and analyzed using a Beckman postcolumn ninhydrin system. Amino-terminal Edman degradations were performed using an Applied Biosystems gas phase sequencer with an in-line PTH analyzer.

Analytical gel filtration and sedimentation

Proteins were chromatographed on a 1 $\times$ 30-cm Superose-12 column equilibrated in T buffer plus 0.5, 1, 2, or 3 M NaCl at 20°C. The flow rate was maintained at 1 ml/min using a Beckman Model 110A pump, and the absorbance at 280 nm of the effluent was monitored using a Kratos Spectroflow 783. For each set of chromatography conditions, elution positions were determined for IgG, bovine serum albumin (BSA), ovalbumin, myoglobin, cytochrome *c*, and vitamin B12. For each run of $\alpha 2$ (reduced or disulfide linked), trace amounts of the standards were mixed with the sample just prior to loading to provide accurate relative elution positions. The experiment of Figure 3 was carried out in 0.5 M NaCl.

The gel filtration experiments were repeated using a 1.5 $\times$ 30-cm Sephacryl column under ambient pressure. Again, the $\alpha 2$ (reduced or disulfide linked) was mixed with the standards and chromatographed through the gel. Positions of the proteins were determined by running SDS gels across the column fractions. The high-performance liquid chromatography (HPLC) and the ambient pressure chromatography gave the same basic results. For reasons we do not understand, the $\alpha 2$ migrated more slowly in both column systems when the salt was raised to 3 M. Thus, in 0.5 M salt, the apparent molecular weights of disulfide-linked and reduced $\alpha 2$ were 90,000 and 35,000, respectively, whereas in 3 M NaCl, these values were 45,000 and 20,000.

For the velocity sedimentation experiment of Figure 4, 5–20% sucrose gradients in T buffer plus 200 mM NaCl were loaded with $\alpha 2$ (disulfide linked or reduced), mixed with a subset of standards used in the HPLC experiments. Gradients (4 ml each) were centrifuged at 45,000 rpm in an SW60 rotor for 16 hr. The contents of the tubes were fractionated, and protein positions determined by running a portion of each fraction on SDS gels.

Operator-binding experiments

Pancreatic DNase I protections were performed in a buffer containing 10 mM Tris-HCl (pH 7), 0.1 mM EDTA, 50 mM NaCl, 5 mM MgCl₂, 2 mM CaCl₂, 2.5 μ g/ml calf-thymus DNA, and 50 μ g/ml BSA. $\alpha 2$ protein or chymotryptic fragments (at concentrations given in the figure legends) and end-labeled operator DNA (at $\sim$ 0.5 nM) were mixed and incubated for 15 min at 20°C, followed by DNase I treatment (0.1 μ g/ml) for 10 min at 20°C. For the values of Table 1, 10-point titration curves were performed, four at a time. For every set, an experiment monitoring the binding of $\alpha 2$ dimers to the wild-type operator was

included as an internal control. Hydroxyl radical and DMS protection experiments were carried out under the same conditions as those described above, using the procedures of Tullius and Dombroski (1986) and Maxam and Gilbert (1977), respectively.

Acknowledgments

We thank Mike Hall, Drew Vershon, Eric Fodor, Bruce Alberts, Kathy Miller, Jack Berry, Pat O'Farrell, Claude Desplan, and Beth Dombroski for valuable experimental advice and discussions, Carl Pabo for pointing out the possible analogy with the *tat* protein, and Bill Werner and Howard Schachman for generously sharing their Model E centrifuge and for performing the equilibrium run. Mike Hall, Cindy Keleher, Caroline Goutte, Drew Vershon, Arkady Mak, Pat O'Farrell, Bruce Alberts, and Ira Herskowitz all contributed valuable comments on the manuscript. We also thank Eric Fodor, Larry Moss, Andreas Meister, and Bill Rutter for their advice and hospitality during our use of their HPLC. This work was supported by grants from the National Institutes of Health and the Pew Scholars Program.

References

- Amann, E., J. Brosius, and M. Ptashne. 1983. Vectors bearing a hybrid *trp-lac* promoter useful for regulated expression of cloned genes in *Escherichia coli*. *Gene* 25: 167–178.
- Astell, C.R., L. Ahlstrom-Jonasson, M. Smith, K. Tatchell, K.A. Nasmyth, and B.D. Hall. 1981. The sequence of the DNAs coding for the mating-type loci of *Saccharomyces cerevisiae*. *Cell* 27: 15–23.
- Ben-Bassat, A. and K. Bauer. 1987. Amino-terminal processing of proteins. *Nature* 326: 315.
- Bender, A. and G.F. Sprague, Jr. 1987. MAT α 1 protein, a yeast transcription activator, binds synergistically with a second protein to a set of cell-type-specific genes. *Cell* 50: 681–691.
- Brake, A., C. Brenner, R. Najarian, P. Laybourn, and J. Merryweather. 1985. Structure of genes encoding precursors of the yeast peptide mating pheromone α -factor. In *Transport and secretion of proteins*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Burkholder, A.C. and L.H. Hartwell. 1985. The yeast α factor receptor: structural properties deduced from the sequence of the STE2 gene. *Nucleic Acids Res.* 13: 8463.
- Desplan, C., J. Theis, and P.H. O'Farrell. 1985. The *Drosophila* developmental gene, engrailed, encodes a sequence-specific DNA-binding activity. *Nature* 318: 630–635.
- Flinta, C., B. Persson, H. Jormvall, and G. Heijne. 1986. Sequence determinants of cytosolic N-terminal protein processing. *Eur. J. Biochem.* 154: 193–196.
- Frankel, A.D., D.S. Bredt, and C.O. Pabo. 1988. Tat protein from human immunodeficiency virus forms a metal-linked dimer. *Science* 240: 70–73.
- Gehring, W.J. 1987. Homeo boxes in the study of development. 1987. *Science* 236: 1245.
- Goutte, C. and A.D. Johnson. 1988. $\alpha 1$ protein alters the DNA-binding specificity of $\alpha 2$ repressor. *Cell* 52: 875–882.
- Hall, M.N. and A.D. Johnson. 1987. Homeo domain of the yeast repressor $\alpha 2$ is a sequence-specific DNA-binding domain but is not sufficient for repression. *Science* 237: 1007–1012.
- Hall, M.N., L. Hereford, and I. Herskowitz. 1984. Targeting of *E. coli* β -galactosidase to the nucleus in yeast. *Cell* 36: 1057–1065.
- Hamer, D.H. 1986. Metallothionein. *Annu. Rev. Biochem.* 55: 913–951.

- Herskowitz, I. and Y. Oshima. 1981. Control of cell type in *Saccharomyces cerevisiae*: Mating type and mating-type inter-conversion. In *Molecular biology of the yeast Saccharomyces: Life cycle and inheritance*. (ed. J.N. Strathern, E.W. Jones, and J.R. Broach) pp. 181–209. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Johnson, A.D. and I. Herskowitz. 1985. A repressor (MAT α 2 product) and its operator control expression of a set of cell type specific genes in yeast. *Cell* **42**: 237–247.
- Keleher, C., C. Goutte, and A. Johnson. 1988. The yeast cell-type specific repressor α 2 acts cooperatively with a non-cell type specific protein. *Cell* (in press).
- Kronstad, J.W., J.A. Holly, and V.L. MacKay. 1987. A yeast operator overlaps an upstream activation site. *Cell* **50**: 369–377.
- Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680–685.
- Laughon, A. and M.P. Scott. 1984. The sequence of a *Drosophila* segmentation gene: Protein structure homology with DNA binding proteins. *Nature* **310**: 25–30.
- Maxam, A. and W. Gilbert. 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci.* **74**: 560–564.
- Miller, A.M., V.L. MacKay, and K.A. Nasmyth. 1985. Identification and comparison of two sequence elements that confer cell-type transcription in yeast. *Nature* **314**: 598–603.
- Nasmyth, K. and D. Shore. 1987. Transcriptional regulation in the yeast life cycle. *Science* **237**: 1162–1170.
- Pabo, C.O. and R.T. Sauer. 1984. Protein-DNA recognition. *Annu. Rev. Biochem.* **53**: 293–321.
- Porter, S.D. and M. Smith. 1986. Homeo-domain homology in yeast MAT α 2 essential for repressor activity. *Nature* **320**: 766–768.
- Ptashne, M. 1986. A genetic switch gene control and phage λ . Cell Press & Blackwell Scientific Publications.
- Shepherd, J.C.W., W. McGinnis, A.E. Carrasco, E.M. DeRobertis, and W.J. Gehring. 1984. Fly and frog homeo domains show homologies with yeast mating type regulatory proteins. *Nature* **310**: 70–71.
- Siebenlist, U., R.B. Simpson, and W. Gilbert. 1980. *E. coli* RNA polymerase interacts homologically with two different promoters. *Cell* **20**: 269–281.
- Sprague, G.F., L.C. Blair, and J. Thormer. 1983. Cell interactions and regulation of cell type in the yeast *Saccharomyces cerevisiae*. *Annu. Rev. Microbiol.* **37**: 623–660.
- Tullius, T.D. and B.A. Dombroski. 1986. Hydroxyl radical footprinting: High resolution information about DNA-protein contracts and application to λ repressor and *cro* protein. *Proc. Natl. Acad. Sci.* **83**: 5469–5473.
- Viera, J. and J. Messing. 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing. *Gene* **19**: 259–268.
- Wilson, K.L. and I. Herskowitz. 1986. Sequences upstream of the STE6 gene required for its expression and regulation by the mating type loci in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.* **83**: 2536–2540.

Flexibility of the yeast alpha 2 repressor enables it to occupy the ends of its operator, leaving the center free.

R T Sauer, D L Smith and A D Johnson

Genes Dev. 1988, **2**:

Access the most recent version at doi:[10.1101/gad.2.7.807](https://doi.org/10.1101/gad.2.7.807)

References

This article cites 28 articles, 7 of which can be accessed free at:
<http://genesdev.cshlp.org/content/2/7/807.full.html#ref-list-1>

License

Email Alerting Service

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