MOLECULAR AND CELLULAR BIOLOGY, Oct. 1990, p. 5077–5086
0270-7306/90/105077-10$02.00/0

Vol. 10, No. 10

Mutations That Define the Optimal Half-Site for Binding Yeast GCN4 Activator Protein and Identify an ATF/CREB-Like Repressor That Recognizes Similar DNA Sites

JOAN W. SELLERS,† ALICE C. VINCENT, AND KEVIN STRUHL*

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115

Received 11 January 1990/Accepted 5 July 1990

The yeast GCN4 transcriptional activator protein binds as a dimer to a dyad-symmetric sequence, indicative of a protein-DNA complex in which two protein monomers interact with adjacent half-sites. However, the optimal GCN4 recognition site, ATGA(C/G)TCAT, is inherently asymmetric because it contains an odd number of base pairs and because mutation of the central C · G base pair strongly reduces specific DNA binding. From this asymmetry, we suggested previously that GCN4 interacts with nonequivalent and possibly overlapping half-sites (ATGAC and ATGAG) that have different affinities. Here, we examine the nature of GCN4 half-sites by creating symmetrical derivatives of the optimal GCN4 binding sequence that delete or insert a single base pair at the center of the site. In vitro, GCN4 bound efficiently to the sequence ATGACGTCAT, whereas it failed to bind to ATGAGCTCAT or ATGATCAT. These observations strongly suggest that (i) GCN4 specifically recognizes the central base pair, (ii) the optimal half-site for GCN4 binding is ATGAC, not ATGAG, and (iii) GCN4 is a surprisingly flexible protein that can accommodate the insertion of a single base pair in the center of its compact binding site. The ATGACGTCAT sequence strongly resembles sites bound by the yeast and mammalian ATF/CREB family of proteins, suggesting that GCN4 and the ATF/CREB proteins recognize similar half-sites but have different spacing requirements. Unexpectedly, in the context of the *his3* promoter, the ATGACGTCAT derivative reduced transcription below the basal level in a GCN4-independent manner, presumably reflecting DNA binding by a distinct ATF/CREB-like repressor protein. In other promoter contexts, however, the same site acted as a weak upstream activating sequence.

GCN4 protein binds to the promoters of many yeast amino acid biosynthetic genes and activates their transcription during conditions of amino acid starvation (2, 18). GCN4 binds as a dimer, and the 60 C-terminal amino acids are sufficient for dimerization and for specific DNA binding (19, 20). The DNA-binding domain shows 45% sequence identity to the Jun oncoprotein (46), and GCN4 and Jun bind similar DNA sites (41). GCN4 contains a leucine zipper motif (25) that is sufficient for dimerization (24, 30, 35) and an adjacent basic region that is involved in specific protein-DNA contacts (1).

The GCN4 recognition sequence has been investigated by saturation mutagenesis of the binding site in the wild-type *his3* promoter (14) and by selection of binding sites from random-sequence DNA (28). Both approaches indicate that a 9-base-pair (bp) dyad-symmetric sequence, ATGA(C/G)TCAT, is optimal for DNA binding and that the central 7 bp are most important. This optimal target resembles the consensus sequence of putative binding sites from 15 GCN4-regulated genes; however none of these naturally occurring sequences match this consensus exactly (14). In the context of the *his3* promoter, the DNA sequence requirements for GCN4 binding and for transcriptional induction in vivo appear indistinguishable (14).

The dyad-symmetric recognition site together with the dimeric nature of GCN4 strongly suggest that the protein-DNA complex consists of two protein monomers interacting with adjacent half-sites. However, the 9-bp GCN4 binding site is very short in comparison with the 15- to 20-bp target sites of many bacterial repressor and activator proteins. Moreover, the crucial positions in the GCN4 site are contiguous and lie within a single turn of the DNA helix. In contrast, the critical residues required for binding by many bacterial repressor and activator proteins are located in two noncontiguous 4- to 5-bp regions in adjacent helical turns of the DNA (3, 5, 6, 9, 26, 45). In these cases, the two monomer–half-site interactions are physically separated by a central region that is not in direct contact with the protein. The precise sequence and dyad-symmetric nature of this central region are relatively unimportant, although the nucleotide composition can indirectly affect the affinity of the interaction by subtle alterations of the DNA structure (23). Thus, the highly compact nature of the GCN4 target sequence strongly suggests that the half-sites recognized by GCN4 monomers are extremely close together and may overlap.

From several lines of evidence, we suggested that GCN4 dimers bind to overlapping and nonequivalent half-sites (14, 20, 28; Fig. 1). First, the optimal binding site contains an odd number of base pairs and hence is inherently asymmetric. This asymmetry is caused by the central C · G base pair; by convention, this position is defined as 0 with the C residue on the top strand and the G residue on the bottom strand. Second, mutation of the central C · G base pair in the target site in the wild-type *his3* promoter to G · C strongly reduces binding. Because the *his3* site differs at one position from the optimal site, such a mutation does not simply reverse the orientation of the sequence (Table 1). Third, mutations at +4 (4 bp downstream of the central base pair) have stronger effects on DNA binding than symmetrical mutations at −4.

* Corresponding author.
† Present address: Transkaryotic Therapies, Inc., Cambridge, MA 02139.

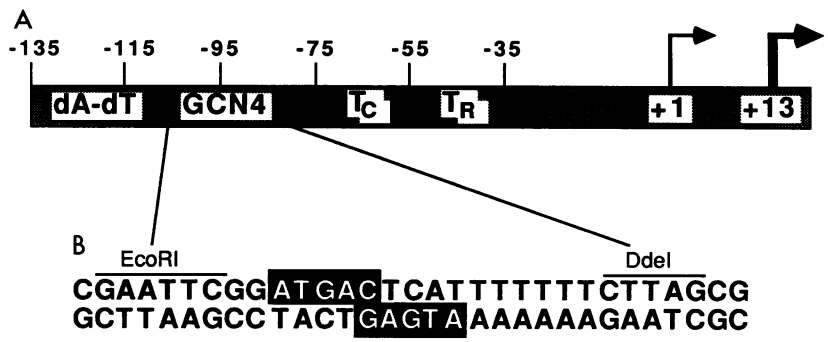

FIG. 1. *his3* promoter region. (A) *his3* region from position −135 to +15, with the promoter elements indicated (drawn to scale). Constitutive *his3* transcription requires the poly(dA-dT) sequence and the T_C element and is initiated with equal efficiency from the +1 and +13 sites. Transcription induced during amino acid starvation requires the GCN4 binding site and the T_R element (TATAAA) and is preferentially initiated from the +13 site (thick arrow). (B) Sequence of the optimal GCN4 binding site flanked by the *Eco*RI and *Dde*I sites used for cloning the degenerate oligonucleotides. The critical bases of the GCN4 binding site (black boxes) depict the half-sites as overlapping at the central C · G base pair.

Fourth, the collection of GCN4 binding sites selected from random-sequence oligonucleotides show distinct sequence preferences at symmetrically equivalent positions. These observations are most consistent with the view that the central C · G base pair is specifically recognized by GCN4 and hence that it is part of both half-sites. However, the possibility that the central C · G base pair indirectly affects DNA binding cannot be excluded.

If the adjacent half-sites in the optimal recognition sequence overlap at the central C · G base pair, they would have distinct DNA sequences, ATGAC and ATGAG (Fig. 1). Moreover, if GCN4 directly interacts with the central C · G base pair, it would certainly prefer one of these half-sites over the other. Here, we determine which of these putative half-sites is optimal for GCN4 binding by analyzing a set of new mutations in the target site in the *his3* promoter. The key feature of these new target sites is that they contain an even number of base pairs and are perfectly symmetric. By removing the asymmetry inherent in GCN4 binding sites containing an odd number of base pairs, the contributions of the individual half-sites can be separated.

One of the symmetrical target sequences examined here, ATGACGTCAT, strongly resembles recognition sites for the ATF/CREB family of DNA-binding proteins that are associated with transcriptional regulation by cyclic AMP or by the adenovirus E1A gene product (12, 32). ATF/CREB-like sites are recognized by a yeast protein(s) and can stimulate transcription when located upstream of a TATA

TABLE 1. DNA-binding and transcriptional activities of GCN4 recognition sequences

Allele (DNA no.)[a]	Sequence[b]	Half-site[c]		Phenotype		RNA[d]	
		Left	Right	In vitro[e]	In vivo[f]	+1	+13
		9-bp sites					
HIS3	ATGACTCTT	ATGAC	AAGAG	++	++	1.0	4.2
his3-Δ88	Deletion			−	−	0.6	0.4
his3-189 (Sc4266)	ATGACTCAT	ATGAC	ATGAG	+++	+++	1.0	30
his3-190 (Sc4272)	ATGAGTCTT	ATGAG	AAGAC	±	±	1.0	1.6
his3-151 (Sc4072)	ATGAATCTT	ATGAA	AAGAT	−	−	1.0[g]	0.7[g]
his3-152 (Sc4074)	ATGATTCTT	ATGAT	AAGAA	−	−	0.9[g]	0.7[g]
his3-192 (Sc4273)	AGGACTCTT	AGGAC	AAGAG	−	−	0.8[g]	0.7[g]
his3-307 (Sc4284)	TTGACTCAA	TTGAC	TTGAG	−	−	1.5	1.1
his3-308 (Sc4285)	GTGACTCAC	GTGAC	GTGAG	++	++	1.0	5.5
his3-309 (Sc4286)	ACGACTCGT	ACGAC	ACGAG	−	−	0.9	0.9
his3-310 (Sc4287)	ATTACTAAT	ATTAC	ATTAG	−	−	0.7	0.9
his3-311 (Sc4288)	ATGTCACAT	ATGTC	ATGTG	−	−	0.5	0.6
		8-bp site					
his3-301 (Sc4280)	ATGATCAT	ATGA	ATGA	−	±	0.5	0.6
		10-bp sites					
his3-303 (Sc4281)	ATGACGTCAT	ATGAC	ATGAC	+	−	0.1	1.3
his3-304 (Sc4282)	ATGAGCTCAT	ATGAG	ATGAG	−	−	0.6	0.8
his3-305 (Sc4283)	ATGATATCAT	ATGAT	ATGAT	−	−	0.5	0.5

[a] The first seven *his3* alleles have been characterized previously (14).
[b] The first nucleotide is at position-100 with respect to the transcriptional initiation site at +1.
[c] Left and Right indicate the two half-sites contained in the GCN4 binding site.
[d] See Materials and Methods and Fig. 2 and 3.
[e] Levels of +1 and +13 RNAs (normalized to the *ded1* RNA level in the same hybridization reaction) when wild-type cells (KY320 derivatives) were grown in 10 mM AT; 1.0 is defined by the level of the +1 RNA in wild-type strains (data from Fig. 4). For each determination, the error in the ratio of +1 to +13 transcripts is approximately ±20%, and the error in the absolute *his3* RNA levels is approximately ±50%.
[f] Determined by growth in 10 mM AT (see Materials and Methods).
[g] Determined previously (42).

element (21, 27). By examining the new GCN4 target sites for their ability to activate *his3* transcription in vivo, we provide evidence for an ATF/CREB-like repressor protein that can compete with GCN4 for binding to the ATGACG TCAT sequence.

MATERIALS AND METHODS

DNA manipulations. All new derivatives of the GCN4 binding site were constructed by using oligonucleotides that were made double stranded by the mutually primed synthesis method (29) and cloned as described previously (14). Briefly, a three-fragment ligation was performed, consisting of a 9-kilobase-pair (kb) *Eco*RI-*Xho*I fragment from YIp55-Sc3384, a 0.9-kb *Dde*I-*Xho*I fragment from pUC8-Sc2676, and the *Eco*RI-*Dde*I-cleaved double-stranded oligonucleotides. The oligonucleotides synthesized were of the following sequences (DNA numbers for the corresponding derivatives are given in parentheses; see Table 1): CGCTA AGAAAAAAAATGATCATCCGAATTCG (Sc4280); CGC TAAGAAAAAAAATGAZZTCATCCGAATTCG (Sc4281, Sc4282, Sc4283); CGCTAAGAAAAAAAXTGAGTCAXC CGAATTCG (Sc4284, Sc4285); CGCTAAGAAAAAAAA XGAGTCXTCCGAATTCG (Sc4286); CGCTAAGAAAAA AAATXAGTXATCCGAATTCG (Sc4287); and CGCTAA GAAAAAAAATGXGXCATCCGAATTCG (Sc4288). Positions noted as Z indicate a mixture of all four nucleotides, and positions noted as X indicate mixtures in which 33% each non-wild-type nucleotide was used. Putative clones for DNAs Sc4280, Sc4281, Sc4282, and Sc4283 were initially screened by digestion with, respectively, *Bcl*I, *Aat*II, *Sac*I, and *Eco*RV (sites created in the desired derivatives); these and all other derivatives were confirmed by DNA sequencing. To facilitate the isolation of fragments for the GCN4 binding reactions, the 1.3-kb *Eco*RI-*Bam*HI fragment containing the entire *his3* gene from these YIp55 derivatives was then subcloned into pUC8. For several derivatives, oligonucleotides derived from *Eco*RI-*Dde*I digestion were blunted and cloned at the blunted *Xho*I site of pRY121Δ20B DNA (47), which lies just upstream of the *gal1* TATA element. The resulting derivatives are identical to those described previously for analysis of ATF/CREB sites in yeast cells (27) except for the precise sequence of the sites.

To obtain oligonucleotide probes for identifying DNA-binding activities from crude extracts, *Eco*RI-*Dde*I fragments containing the symmetrical target sequences of Sc4281 and Sc4282 were blunted and cloned into the *Sma*I site of pBluescript-KS (Stratagene). The resulting molecules were cleaved with *Eco*RI and *Bam*HI and labeled at their 5′ ends with ^{32}P, and the 50-bp fragments containing the symmetrical target sites were isolated by electrophoresis in acrylamide gels.

GCN4 binding studies. In vitro synthesis of ^{35}S-labeled GCN4 and analysis of DNA-protein complexes were carried out essentially as described previously (14, 18) except that in some cases poly(dI-dC) at a final concentration of 100 μg/ml was used in the binding buffer instead of salmon sperm DNA. For the competition binding experiments, GCN4 was incubated with a mixture containing 15 nM a 631-bp *Pvu*II-*Rsa*I fragment derived from the pUC8 subclones described above and 7 nM the 407-bp fragment obtained by *Nde*I cleavage of the 631-bp fragment from the wild-type *his3* promoter. The relative affinities of the target sequences are directly related to band intensities of the protein-DNA complexes (14, 18) and are defined as follows: more avid than the site in the wild-type *his3* promoter (+++); equivalent to the wild-type site (++); less avid than the wild-type site (+); binds poorly (±); binding not detectable (−).

Identification of DNA-binding activities from yeast cell extracts. A crude yeast extract was prepared from a derivative of strain KY320 (8) by a modified version of a previously described procedure (48). Briefly, spheroplasts prepared from cells grown to mid-log phase in YPD (conditions preventing GCN4 synthesis) (15, 43) were disrupted by Dounce homogenization. The extract was made to 0.3 M ammonium sulfate and cleared by centrifugation; the resulting protein was precipitated with 2 volumes of saturated ammonium sulfate and then suspended in buffer (20 mM *N* - 2 - hydroxyethylpiperazine - *N*′ - 2 - ethanesulfonic acid [HEPES]-NaOH [pH 7.5], 10% glycerol, 1 mM EDTA, 1 mM dithiothreitol, 0.1 mM phenylmethylsulfonyl fluoride, 0.4 μg of pepstatin A per ml, 1 μg of chymostatin per ml, 2.5 μg of antipain per ml, 0.2 μg of aprotinin per ml, 0.1 mM benzamidine, 0.5 μg of leupeptin per ml) to give a final salt concentration of 100 mM; the final protein concentration was 1.8 mg/ml. The protein was chromatographed on a heparin-agarose column at a protein/resin ratio of 33 mg/ml and then eluted with a 100 mM to 1 M gradient of KCl. The resulting fractions were combined with 1 ng of ^{32}P-labeled oligonucleotide probes in 15 μl of buffer containing 20 mM Tris hydrochloride (pH 7.9), 50 mM KCl, 1 mM EDTA, 1 mM dithiothreitol, 100 μg of bovine serum albumin per ml, and 160 μg of poly(dI-dC) per ml and incubated at room temperature for 15 min. After addition of 5 μl of 20% glycerol, the products were separated on an 8% native polyacrylamide gel containing 90 mM Tris-borate (pH 8.3).

Phenotypic analysis. YIp55 DNA molecules containing the various GCN4 binding sites were introduced into yeast cells by precise replacement of the *his3* chromosomal location as described previously (8, 14, 39). The recipient yeast strains were KY320 (relevant genotype *ura3-52 his3-Δ200*) (8), KY329 (relevant genotype *ura3-52 his3-TRP1 gcn4-Δ1*) (41), and KY603 (relevant genotype *ura3-52 his3-Δ200 gcd1-1*) (14). As an initial phenotypic test, KY320 derivatives were examined for growth in the presence of aminotriazole (AT), a competitive inhibitor of the *his3* gene product. As shown previously, the degree of AT resistance is directly related to the level of *his3* RNA (8, 14, 41).

To measure constitutive *his3* RNA levels directly, KY320 or KY329 derivatives were grown at 30°C in YPD broth. To measure RNA levels under inducing conditions, KY603 derivatives were grown in YPD broth at 23°C (because of the temperature-sensitive *gcd1-1* mutation) or KY320 derivatives were grown in glucose minimal medium containing 10 mM AT. RNA was hybridized to completion with an excess of ^{32}P-end-labeled *his3* and *ded1* oligonucleotide probes and treated with S1 nuclease, and the products were analyzed on 10% denaturing polyacrylamide gels (8). The levels of *his3* +1 and +13 RNAs were quantitated by densitometry of appropriately exposed autoradiograms and normalized to the level of the *ded1* internal control. For each determination, the error in the ratio of +1 to +13 RNAs is ±20%, and the error in the absolute *his3* RNA levels is ±50%.

The transcriptional activity of promoters containing the symmetrical target sequences upstream of the *gal1* TATA element was determined by introducing the relevant DNA molecules into KY320. Cells were grown in minimal medium lacking uracil (to maintain the unstable high-copy-number plasmids) and assayed for β-galactosidase activity as described previously (19). Enzyme activities were determined in triplicate and normalized to the optical density at 600 nm of the cultures; the values are accurate to ±30%.

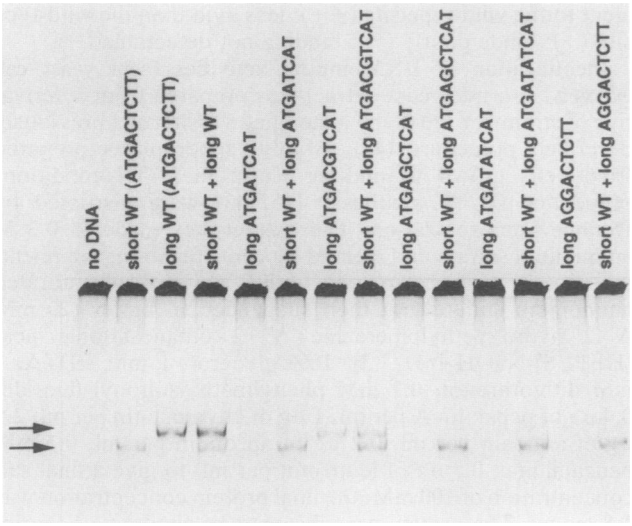

FIG. 2. GCN4 binding to 8- and 10-bp sites. In vitro-synthesized ^{35}S-labeled GCN4 protein was incubated with short (407-bp *Pvu*II-*Nde*I) and/or long (631-bp *Pvu*II-*Rsa*I) DNA fragments containing the indicated binding sites. Positions corresponding to protein complexes with the long (upper arrow) and short (lower arrow) DNA fragments are indicated on the left. WT, Wild type.

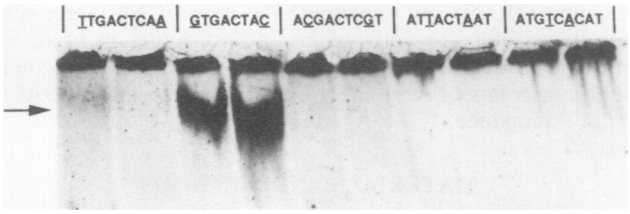

FIG. 3. GCN4 binding to symmetrically mutated 9-bp sites. In vitro-synthesized ^{35}S-labeled GCN4 protein was incubated with the 631-bp *Pvu*II-*Rsa*I fragment containing the indicated site. Alternate lanes contain, respectively, salmon sperm DNA or poly(dI-dC) in the binding buffer. The position corresponding to the protein-DNA complexes is indicated on the left.

RESULTS

GCN4 binding to symmetrical target sequences. On the basis of a variety of observations (summarized in the introduction), we proposed that GCN4 dimers interact with nonequivalent half-sites (ATGAC and ATGAG) that have different affinities (14, 20, 28). To distinguish the contributions of the individual half-sites, symmetrical target sequences consisting of two copies of a particular half-site were generated (Table 1). Three of these symmetrical targets contain an additional base pair at the center of the GCN4 site and correspond to pairs of ATGAC, ATGAG, or ATGAT half-sites. Another symmetrical target was created by deleting the central C · G base pair to produce a pair of ATGA half-sites.

The relative affinities of these symmetrical target sequences were determined by competition binding assays (Fig. 2). Specifically, ^{35}S-labeled GCN4 protein was synthesized in vitro and incubated with a mixture of a 631-bp fragment containing the target site of interest and a 407-bp fragment containing the GCN4 binding site in the wild-type *his3* promoter. Since the electrophoretic mobility of protein-DNA complexes in native acrylamide gels is affected by the molecular weight of the DNA, the complexes with the wild-type *his3* target have a faster mobility than complexes with the mutant derivatives and hence can be easily distinguished. Under conditions of the assay, the intensities of bands corresponding to the protein-DNA complexes are directly related to the binding affinity (14, 18).

The results of these experiments indicate that GCN4 bound to the DNA fragment containing the sequence AT GACGTCAT with nearly equal efficiency as it did to the site in the wild-type *his3* promoter (Fig. 2; Table 1). However, GCN4 bound extremely poorly to all other sites tested; indeed, these derivatives behaved indistinguishably from a 9-bp target site containing the point mutation AGGAC TCTT. Thus, GCN4 binds to a target containing a pair of ATGAC half-sites but not to targets containing pairs of ATGAG, ATGAT, or ATGA half-sites. In addition, we examined derivatives of the optimal 9-bp target sequence containing symmetrically disposed mutations (Fig. 3; Table 1). Of these, GCN4 bound only to GTGACTCAC, a sequence equivalent to that found in the high-affinity target site in the *his4* promoter (2) and containing deviations from the optimal binding site that are tolerated (14, 28). Taken together, these results confirm the importance of the central C · G base pair for binding by GCN4 and suggest that the optimal half-site is ATGAC. Furthermore, they indicate that despite the compact nature of the GCN4 recognition sequence, the protein can accommodate the insertion, but not a deletion, of a single base pair in the center of its binding site.

ATGACGTCAT binds GCN4 in vitro but is defective for transcriptional activation in vivo. To determine how these mutant sites function in vivo, DNA molecules containing the symmetrical target sequences but otherwise identical to the wild-type *his3* promoter (Fig. 1) were introduced in a single copy at the *his3* locus by gene replacement. The resulting strains were then tested for growth in the presence of AT, a competitive inhibitor of the *his3* gene product (Table 1). As determined previously, the degree of AT resistance is directly related to the level of *his3* transcription (8, 14, 41).

Saturation mutagenesis of the GCN4 binding site in the *his3* promoter indicates that the DNA sequence requirements for DNA binding in vitro and transcriptional activation in vivo are indistinguishable (14). Thus, it is not surprising that the *his3* alleles containing target sites that do not detectably interact with GCN4 grow extremely poorly in the presence of AT. On the other hand, the strain containing *his3-308*, which contains a functional GCN4 binding site similar to that found in the *his4* promoter (2), grows well under these conditions. Surprisingly, however, cells containing the *his3-303* allele (ATGACGTCAT) grow slowly in the presence of AT despite the fact that GCN4 binds efficiently to this sequence in vitro.

One explanation for this unexpected result is that *his3-303* is an unusual allele that permits GCN4 to bind but not to activate. Alternatively, the GCN4 target sequence in *his3-303* might also be recognized by a distinct protein that prevents GCN4 from binding and activating. To distinguish between these possibilities, *his3* RNA levels under a variety of experimental conditions were quantitated by S1 analysis, using the level of *ded1* RNA as an internal control.

ATGACGTCAT causes transcriptional repression that is independent of GCN4 protein. Under normal growth conditions, the *his3* gene is transcribed with equal effficiency from two distinct initiation sites defined as +1 and +13 (Fig. 1). This basal level of transcription is not affected by point mutations or small deletions in the GCN4 binding site (14, 42). However, when *his3* transcription is induced by GCN4

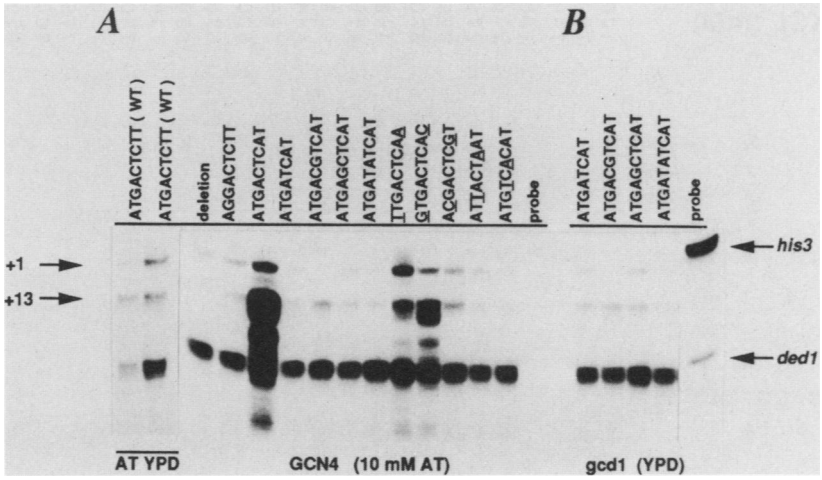

FIG. 4. *his3* RNA levels under inducing conditions. (A) Analysis of RNA from KY320 derivatives containing the indicated target site or from the isogenic *HIS3* wild-type (WT) strain KY114 (leftmost two lanes) that were grown in YPD or in minimal medium containing 10 mM AT. The lane marked deletion indicates a strain containing *his3-Δ88*; the twofold decrease of the +13 transcript probably reflects the loss of activation mediated through the GCN4 binding site that appears to depend on a distinct factor (42). The additional band observed in ATGACTCAT and GTGACTCAC strains is the +22 transcript that is fully inducible by GCN4 but inefficiently used (14, 40, 42). (B) Analysis of RNA from KY603 derivatives (*gcd1-1*) containing the indicated GCN4 binding sites that were grown in YPD. Positions of the +1 and +13 *his3* RNAs are indicated on the left; positions of the untreated *his3* and *ded1* probes are indicated on the right. Quantitative analysis of these data is presented in Tables 1 and 2.

under conditions of amino acid starvation or in *gcd1-1* strains, initiation from the +13 site is increased fivefold, whereas initiation from the +1 site is unchanged (40).

In accord with the growth properties of strains containing any of the mutant alleles, no increase in the *his3* +13 transcript was observed under circumstances in which transcriptional induction by GCN4 should occur (Fig. 4; Tables 1 and 2). These include normal (*GCD1*$^+$) cells grown in the presence of AT (standard inducing conditions due to histidine limitation) and *gcd1-1* cells grown in nonstarvation conditions (which results in the constitutive synthesis of GCN4) (15, 43). Interestingly, under either circumstance, the +1 transcript in strains containing the *his3-303* allele was decreased by a factor of 5 to 10 below basal levels, as if transcriptional repression were occurring.

To determine whether this apparent repression was independent of GCN4 protein, RNA was prepared from noninduced cells, in which the very low levels of GCN4 protein are insufficient for induction (15, 43), and from cells containing a GCN4 deletion (Fig. 5; Table 2). Strikingly, *his3* transcription from *his3-303* was almost completely eliminated under either circumstance. In contrast, transcription from the wild-type or any of the other mutant alleles occurred at the normal basal level. These results indicate that the ATGACGTCAT target sequence in *his3-303* leads to transcriptional repression in a GCN4-independent manner.

TABLE 2. Transcriptional activities in wild-type, *gcd1*, and *gcn4* strains

Allele	Sequence[a]	*his3* RNA levels (+1, +13 transcripts)[b]			
		gcd1	Wild type		*gcn4*
			AT	YPD	
HIS3	ATGACTCTT	1.0, 5.0[c]	1.0, 4.2	1.0, 1.0[c]	0.7, 0.4
his3-301	ATGATCAT	0.6, 0.9	0.5, 0.6	1.5, 1.2	0.9, 1.0
his3-303	ATGACGTCAT	0.1, 1.2	0.1, 1.3	<0.1, 0.1	<0.1, 0.1
his3-304	ATGAGCTCAT	0.6, 0.8	0.6, 0.8	1.1, 1.4	1.0, 0.7
his3-305	ATGATATCAT	0.5, 0.7	0.5, 0.5	0.7, 0.5	0.9, 0.7

[a] The first nucleotide is at position −100 with respect to the transcriptional initiation site at +1.

[b] Levels of +1 and +13 RNAs (normalized to the *ded1* RNA level in the same hybridization reaction) in *gcd1* strains (KY603 derivatives), wild-type strains (KY320 derivatives grown in YPD or 10 mM AT medium), and *gcn4* strains (KY329 derivatives); data from Fig. 4 and 5. The level of the wild-type *his3* gene in the wild-type strain is defined as 1.0. The error in the ratio of +1 to +13 transcripts is approximately ±20%, and the error in the absolute *his3* RNA levels is approximately ±50%.

[c] Determined previously (14, 42).

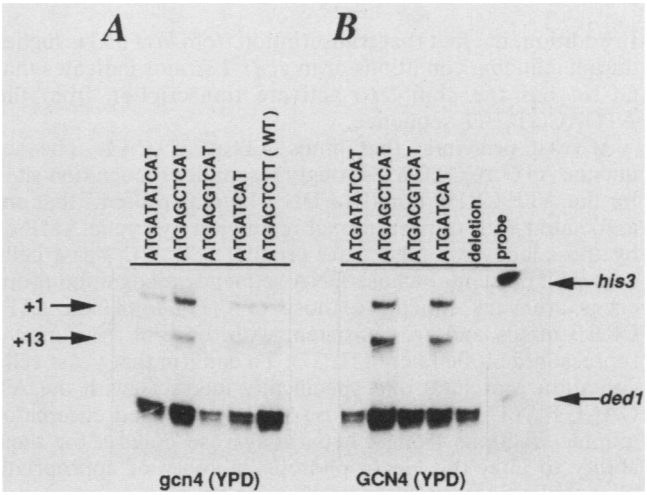

FIG. 5. *his3* RNA levels under noninducing conditions. Shown is analysis of RNA from KY803 (*gcn4-Δ1*) derivatives (A) or KY320 derivatives (B) containing the indicated target site that were grown in YPD; the lane marked deletion indicates a strain containing *his3-Δ200*, which deletes the entire *his3* promoter region and structural gene. The positions of the +1 and +13 *his3* RNAs and the *ded1* RNA are indicated. Quantitative analysis of these data is presented in Table 2. WT, Wild type.

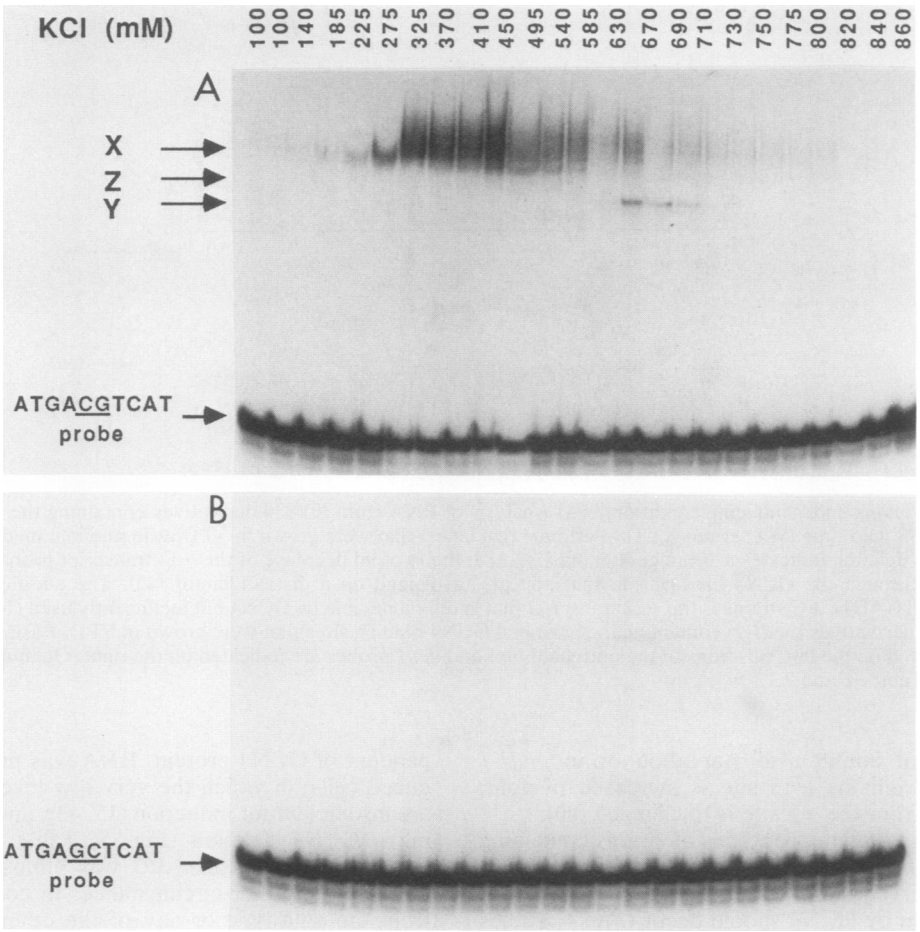

FIG. 6. Yeast proteins binding to symmetrical target sequences. Chromatographic fractions of a yeast extract eluted from a heparin-agarose column by the indicated concentrations of KCl were incubated with oligonucleotides containing ATGACGTCAT (A) or ATGAGCTCAT (B). Positions of bands corresponding to unbound probe and complexes X, Y, and Z are indicated. The complexes are GCN4-independent because the extract was prepared from cells grown in YPD, conditions that prevent the synthesis of GCN4 (15, 43).

In addition, the fact that transcription from *his3-303* is higher during inducing conditions or in *gcd1-1* strains indicates that GCN4 has the ability to activate transcription from the ATGACGTCAT sequence.

A yeast protein(s) that binds ATGACGTCAT. The sequence ATGACGTCAT strongly resembles recognition sites for the ATF/CREB family of DNA-binding proteins that are associated with transcriptional regulation by cyclic AMP or by the adenovirus E1A gene product (12, 32). Yeast cells contain a protein(s) whose DNA sequence recognition properties are very similar to those of the mammalian ATF/CREB family and are consistent with those of the putative repressor identified here (21, 27). To confirm that yeast cells contain a protein(s) that specifically interacts with the ATGACGTCAT sequence in *his3-303*, we assayed chromatographic fractions from a heparin-agarose column for their ability to alter the electrophoretic mobility of appropriate oligonucleotides in native acrylamide gels.

At least two, and possibly three, chromatographically distinct activities that bind the ATGACGTCAT oligonucleotide were defined by their ability to form electrophoretically distinct protein-DNA complexes (Fig. 6). Complex X was observed primarily with fractions eluting between 275 and 410 mM KCl, complex Y (barely visible in the exposure shown in Fig. 6) was observed with fractions eluting between 450 and 540 mM KCl, and complex Z was observed with fractions eluting between 630 and 690 mM KCl. Complexes X, Z, and probably Y represent interaction with specific DNA-binding factors because they are not detected in parallel assays using the ATGAGCTCAT oligonucleotide. Like other ATF/CREB proteins, the factor(s) described here binds poorly to the optimal GCN4 binding site (data not shown). At this stage of the analysis, we cannot determine whether the multiple activities represent distinct protein factors or differentially modified or proteolytic products of a single protein. In addition, it is likely but difficult to prove that the factors described here are the same as previously described yeast ATF/CREB-like factors that were identified by binding to very similar, but not identical, DNA sequences (21, 27).

ATGACGTCAT weakly activates transcription in other promoter contexts. In apparent contrast to the results in Fig. 4 and 5, previous studies indicate that sequences recognized by yeast ATF/CREB protein activate transcription when fused to the *cyc1* or *gal1* TATA element (21, 27). Since the ATF/CREB sites used in these earlier studies differed slightly from that present in *his3-303*, we analyzed the transcriptional activity of promoters containing the symmetrical target sequences upstream of the *gal1* TATA element and mRNA initiation region. The constructs are identical to

TABLE 3. Activities of promoters containing target sites upstream of the *gal1* TATA element

Plasmid DNA	Sequence upstream of *gal1* TATA element[a]	β-Galactosidase activity[b]
ΔUASG/βG[c]	None	0.2
ATF1G/βG[c]	gACGTCATccc	5.6
Sc4289	aattccggATGACGTCATttttttttctt	6.6
Sc4290	aattccggATGAGCTCATttttttttctt	0.7
Sc4291	(aattccggATGACGTCATttttttttctt)$_2$	29.7
Sc4292	(aattccggATGAGCTCATttttttttctt)$_2$	2.5
LR1Δ20B[d]	UAS_{gal}	900

[a] Oligonucleotides inserted in one or two copies at the *Xho*I site of the multicopy plasmid ΔUASG/βG (27), which lies just upstream of the *gal1* TATA element in a *gal1-lacZ* fusion gene. The sequences present in Sc4289 and Sc4291 were derived from the *his3-303* allele; the sequences in Sc4290 and Sc4292 were derived from the *his3-304* allele.

[b] Cells grown in glucose minimal medium lacking uracil except for those harboring LR1Δ20B, which were grown in galactose minimal medium. Activities (Miller units per 10^7 cells) represent the average of three independent transformants; the error is approximately ±30%.

[c] Described previously (27).

[d] Described previously (47).

those analyzed previously (27) except for the precise sequence of the target sites.

One or two copies of the ATGACGTCAT sequence in *his3-303* clearly activated transcription above the level observed in the control plasmid, which lacks any upstream promoter element (ΔUASG/βG) (Table 3). This activation was specific to the ATGACGTCAT sequence because expression was 10-fold lower in analogous constructs containing the ATGAGCTCAT sequence from *his3-304*. Activation mediated by ATGACGTCAT was very weak; two copies of this sequence resulted in only about 3% of the level achieved by UAS_G, which corresponds roughly to 20% of the basal level of *his3* transcription. The extremely low level of activation conferred by a single copy of ATGACGTCAT appears to contradict previous results which indicated that the single ATF site present in plasmid ATF1G/βG strongly stimulated transcription (27). However, we find that plasmid ATF1G/βG confers the same low level of expression as that observed with Sc4289, which contains a single copy of the ATF/CREB site from *his3-303*. This discrepancy does not reflect differences in yeast strains; results identical to those shown in Table 3 were obtained when the plasmids were introduced into YSC294, the strain used in the previous study (27). Thus, the ATGACGTCAT sequence in *his3-303* can weakly stimulate transcription when located just upstream of the *gal1* TATA element.

DISCUSSION

Nature of the GCN4 recognition sequence. Although the GCN4-DNA complex consists of two protein monomers interacting with adjacent half-sites, the optimal recognition site, ATGA(C/G)TCAT, is inherently asymmetric because the central C · G base pair is functionally important (14, 20, 28). In principle, the central base pair could contribute to DNA-binding affinity either directly by a specific interaction with GCN4 or indirectly as a result of the inherent DNA structure of the target site. The results presented here strongly argue against an indirect mechanism such as the DNA flexibility that influences the affinity of bacteriophage 434 repressor for its operator (23). In that case, the central region not contacted by the protein extends for 4 to 6 bp, and its effect on affinity depends on the overall nucleotide composition, not the specific sequence. In contrast, GCN4 efficiently recognizes ATGACGTCAT but not ATGAGCT CAT even though both sequences have the same nucleotide composition. Moreover, GCN4 can efficiently recognize both 9- and 10-bp target sequences that clearly have very different DNA structures. Thus, the high sequence specificity in the central region of both the 9- and 10-bp GCN4 recognition sites strongly support a direct interaction model.

Assuming that GCN4 directly contacts the central region, the DNA-binding properties of the symmetrical target sequences indicate that the optimal half-site is ATGAC, not ATGAG. This would mean that the adjacent half-sites in the 9-bp optimal recognition sequence are nonequivalent because of their overlap at the central C · G base pair. A further implication is that when GCN4 interacts with the optimal 9-bp target sequence, the left half-site (ATGAC) contributes more to the overall affinity than the right half-site (ATGAG). Presumably, one monomer of the GCN4 dimer contacts the left-half site and the central base pair, whereas the monomer interacting with the right half-site does not contact the central position. This view of the GCN4-DNA interaction (Fig. 7) accounts for why alterations in the right half-site are tolerated better than symmetrically equivalent alterations in the left half-site (28).

GCN4 appears to be a surprisingly flexible protein because it can accommodate a major structural disruption, the insertion of a single base pair, in the center of its otherwise compact binding site (Fig. 7). In part, this accommodation reflects the fact that ATGACGTCAT contains two optimal half-sites; however, GCN4 clearly prefers a 9-bp target despite the impossibility of both half-sites being optimal. Although many DNA-binding proteins are highly sensitive to spacing changes in the target site (5, 6, 45), some proteins tolerate or even prefer different spacings between half-sites (9, 33, 34, 37). However, in all of these cases of flexibility, the precise sequence at the center of the binding site is relatively unimportant, and the protein dimerization region resides in a structural domain distinct from the region required for DNA contacts. Thus, one could imagine that the DNA interaction surfaces of the two monomers are relatively independent. In contrast, the dimerization and DNA-binding functions of GCN4 are localized to the 60 C-terminal residues (19, 20), a region that appears to be a single structural domain as determined by proteolytic mapping (17).

The most likely structural basis for the surprising flexibility of GCN4 is a hinge region between the leucine zipper that is necessary and sufficient for dimerization (24, 30, 35) and the adjacent basic region that directly contacts the DNA (1). However, this explanation requires a very short hinge (less than 10 amino acids) and does not account for the highly conserved spacing relationship between the leucine zipper and adjacent basic region observed in leucine zipper proteins (25). A model summarizing all of these considerations is shown in Fig. 7.

Relationship between GCN4 and ATF/CREB recognition sequences. Members of the mammalian ATF/CREB family of proteins are defined by their ability to bind sites that resemble the dyad-symmetric sequence TGACGTCA (12, 32). Like GCN4, these proteins bind as dimers to their target sites, and they contain leucine zipper motifs and adjacent basic regions (11, 13, 16); thus, it is extremely likely that the ATF/CREB family recognizes adjacent TGAC half-sites. The putative ATF/CREB half-site corresponds precisely to the central region of the optimal GCN4 half-site determined here. This suggests that GCN4 and the ATF/CREB protein

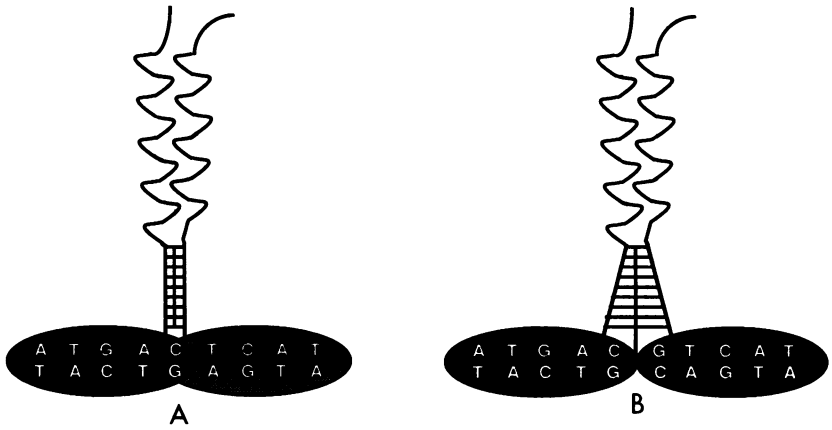

FIG. 7. Model for GCN4 binding. The subregions of the GCN4 DNA-binding domain are depicted as follows: leucine zipper dimerization surface (parallel wavy lines); hinge or spacer region (cross-hatch); DNA interaction surface (shaded ovals, each representing a monomer contacting residues within the half-site). Binding to the optimal 9-bp target site (A) involves nonequivalent monomer–half-site interactions, with the left monomer (black oval) contributing more to the overall affinity than the right monomer (gray oval). Binding to the optimal 10-bp target site (B) involves equivalent and optimal monomer–half-site interactions but necessitates structural deformity of the region between the leucine zipper and DNA-binding surface.

family recognize similar half-sites but differ in their requirements regarding the spacing between half-sites. In support of this idea, the mammalian AP-1 protein family, which recognizes the same sequences as GCN4 (7, 41), is immunologically related to the ATF/CREB protein family (12). Thus, we propose that the GCN4/AP-1 and ATF/CREB classes of proteins belong to the same evolutionarily conserved superfamily of proteins that recognize essentially identical half-sites. A precedent for members of a protein superfamily that recognize similar half-sites with distinct spatial constraints has been suggested to explain the DNA-binding properties of the estrogen and thyroid hormone receptors (10, 44).

Evidence for an ATF/CREB-like repressor that competes with GCN4 for binding ATGACGTCAT. Constitutive *his3* transcription requires a poly(dA-dT) sequence (−130 to −115) and the T_C downstream element (−80 to −55) but is not affected by GCN4 protein (38, 40). Moreover, numerous mutations that inactivate or delete the GCN4 binding site, including several described in this report, do not influence the basal level of *his3* expression (14, 42). In contrast, the level of constitutive *his3* transcription is significantly reduced in *his3-303* strains in a GCN4-independent manner. This suggests that the particular sequence of this allele, ATGACGTCAT, is the target of a transcriptional repressor. Presumably, this putative repressor does not recognize the GCN4 binding site in the wild-type *his3* promoter or any of the other mutated sites and hence has relatively specific DNA-binding properties. In accord with this hypothesis, yeast cells contain at least one, and possibly several, proteins that bind to the target site in the *his3-303* allele. Moreover, these activities are unable to bind ATGAGCT CAT (Fig. 6) or the optimal GCN4 binding site (data not shown) and hence display the binding specificity expected for the putative repressor. However, there is no direct evidence that implicates any of these observed DNA-binding activities as being required for transcriptional repression in vivo.

The fact that GCN4 protein induces transcription from *his3-303* above the repressed level indicates that GCN4 has the ability to activate transcription when bound to the ATGACGTCAT sequence. In this circumstance, cell growth is strongly inhibited by AT, and the absolute level of *his3* RNA is only about 25% of the wild-type induced level, even though GCN4 binds with similar affinity to ATGACGTCAT and the *his3* wild-type sequence. The simplest interpretation of this observation is that GCN4 and the putative repressor compete for binding to ATGACGTCAT in vivo.

The sequence ATGACGTCAT strongly resembles recognition sites for the ATF/CREB family of DNA-binding proteins that are associated with transcriptional regulation by cyclic AMP or by the adenovirus E1A gene product (12, 32). As shown here and elsewhere (21, 27), yeast cells contain a protein(s) whose DNA sequence recognition properties are very similar to those of the mammalian ATF/CREB family and are consistent with those of the putative repressor identified here. Thus, it is extremely likely that the repression observed in the *his3-303* promoter is due to an ATF/CREB-like protein.

Function of the ATF/CREB site in yeast cells depends on promoter context. Although the ATGACGTCAT sequence in *his3-303* clearly represses transcription in a GCN4-independent manner, previous studies indicate that related sequences recognized by a yeast ATF/CREB protein(s) activate transcription when fused to the *CYC1* TATA element (21, 27). This apparent difference does not reflect subtle distinctions between the ATF/CREB sites, because the sequence present in *his3-303* can weakly activate transcription when located upstream of the *gal1* TATA element (Table 3). Thus, it is clear that the ATGACGTCAT sequence in *his3-303* can act positively or negatively, depending on the promoter context.

One explanation for this apparent discrepancy is that a single yeast ATF/CREB protein serves as a transcriptional activator or repressor, depending on the promoter context. Such a mechanism has been suggested for both the yeast MCM1 (4, 22, 31) and RAP1 (36) proteins. Alternatively, the discrepancy might reflect the existence of distinct activator and repressor proteins that have similar (but possibily nonidentical) DNA-binding specificities; indeed, the experiments here suggest the presence of distinct ATF/CREB binding activities. By either of these models, it is presumed that the activator and repressor act differently in combination with other promoter elements, presumably reflecting distinct protein-protein interactions. With respect to the second model, the apparent activation mediated by the *his3-303* sequence (and the other ATF/CREB sites) might

not reflect promoter context effects but rather might be due to titrating out the repressor protein, especially since all of the relevant promoters are present on high-copy-number plasmids. The resolution of these issues awaits the identification of the gene(s) encoding yeast ATF/CREB protein(s) and the phenotypic consequences in strains lacking such a protein(s).

ACKNOWLEDGMENTS

We thank Rick Wobbe for improving the DNA-binding assay and for help in the preparation of yeast cell extracts and RNA quantitations and Young-Sun Lin and Michael Green for plasmids and strains.

This work was supported by a Public Health Service postdoctoral fellowship to J.W.S. from the National Institutes of Health and by grants to K.S. from the National Institutes of Health (Public Health Service grant GM 30186) and from the Lucille Markey Trust.

LITERATURE CITED

1. **Agre, P., P. F. Johnson, and S. L. McKnight.** 1989. Cognate DNA binding specificity retained after leucine zipper exchange between GCN4 and C/EBP. Science **246**:922–926.
2. **Arndt, K., and G. Fink.** 1986. GCN4 protein, a positive transcription factor in yeast, binds general control promoters at all 5′ TGACTC 3′ sequences. Proc. Natl. Acad. Sci. USA **83**:8516–8520.
3. **Bass, S., P. Sugiono, D. N. Arvidson, R. P. Gunsalus, and P. Youderian.** 1987. DNA specificity determinants of *Escherichia coli* tryptophan repressor binding. Genes Dev. **1**:565–572.
4. **Bender, A., and G. F. Sprague Jr.** 1987. MATα1 protein, a yeast transcription activator, binds synergistically with a second protein to a set of cell-type-specific genes. Cell **50**:681–691.
5. **Benson, N., P. Sugiono, and P. Youderian.** 1988. DNA sequence determinants of λ repressor binding *in vivo*. Genetics **118**:21–29.
6. **Benson, N., and P. Youderian.** 1989. Phage λ Cro protein and *cI* repressor use two different patterns of specific protein-DNA interactions to achieve sequence specificity *in vivo*. Genetics **121**:5–12.
7. **Bohmann, D., T. J. Bos, A. Admon, T. Nishimura, P. K. Vogt, and R. Tjian.** 1987. Human proto-oncogene c-*jun* encodes a DNA binding protein with structural and functional properties of transcription factor AP-1. Science **238**:1386–1392.
8. **Chen, W., and K. Struhl.** 1988. Saturation mutagenesis of a yeast *his3* TATA element: genetic evidence for a specific TATA-binding protein. Proc. Natl. Acad. Sci. USA **85**:2691–2695.
9. **Falvey, E., and N. D. F. Grindley.** 1987. Contacts between γδ resolvase and the γδ *res* site. EMBO J. **6**:815–821.
10. **Glass, C. K., J. M. Holloway, O. V. Devary, and M. G. Rosenfeld.** 1988. The thyroid hormone receptor binds with opposite transcriptional effects to a common sequence motif in thyroid hormone and estrogen response elements. Cell **54**:313–323.
11. **Gonzalez, G. A., K. K. Yamamoto, W. H. Fischer, D. Karr, P. Menzel, W. Biggs, W. W. Vale, and M. R. Montminy.** 1989. A cluster of phosphorylation sites on the cyclic-AMP regulated nuclear factor CREB predicted by its sequence. Nature (London) **337**:749–752.
12. **Hai, T., F. Liu, E. A. Allegretto, M. Karin, and M. R. Green.** 1988. A family of immunologically related transcription factors that includes multiple forms of ATF and AP-1. Genes Dev. **2**:1216–1226.
13. **Hai, T., F. Liu, W. J. Coukos, and M. R. Green.** 1989. Transcription factor ATF cDNA clones: an extensive family of leucine zipper proteins able to selectively form DNA-binding heterodimers. Genes Dev. **3**:2083–2090.
14. **Hill, D. E., I. A. Hope, J. P. Macke, and K. Struhl.** 1986. Saturation mutagenesis of the yeast *HIS3* regulatory site: requirements for transcriptional induction and for binding by GCN4 activator protein. Science **234**:451–457.
15. **Hinnebusch, A. G.** 1984. Evidence for translational regulation of the activator of general amino acid control in yeast. Proc. Natl. Acad. Sci. USA **81**:6442–6446.
16. **Hoeffler, J. P., T. E. Meyer, Y. Yun, J. L. Jameson, and J. F. Haebner.** 1988. Cyclic AMP-responsive DNA-binding protein: structure based on a cloned placental cDNA. Science **242**:1430–1433.
17. **Hope, I. A., S. Mahadevan, and K. Struhl.** 1988. Structural and functional characterization of the short acidic transcriptional activation region of yeast GCN4 protein. Nature (London) **333**:635–640.
18. **Hope, I. A., and K. Struhl.** 1985. GCN4 protein, synthesized in vitro, binds to *HIS3* regulatory sequences: implications for the general control of amino acid biosynthetic genes in yeast. Cell **43**:177–188.
19. **Hope, I. A., and K. Struhl.** 1986. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. Cell **46**:885–894.
20. **Hope, I. A., and K. Struhl.** 1987. GCN4, a eukaryotic transcriptional activator protein, binds as a dimer to target DNA. EMBO J. **6**:2781–2784.
21. **Jones, R. H., and N. C. Jones.** 1989. Mammalian cAMP-responsive element can activate transcription in yeast and binds a yeast factor(s) that resembles the mammalian transcription ATF. Proc. Natl. Acad. Sci. USA **86**:2176–2180.
22. **Keleher, C. A., C. Goutte, and A. D. Johnson.** 1988. The yeast cell-type-specific repressor α2 acts cooperatively with a non-cell-type-specific protein. Cell **53**:927–936.
23. **Koudelka, G. B., P. A. B. Harbury, S. Harrison, and M. Ptashne.** 1988. DNA twisting and the affinity of bacteriophage 434 operator for bacteriophage 434 repressor. Proc. Natl. Acad. Sci. USA **85**:4633–4637.
24. **Kouzarides, T., and E. Ziff.** 1989. Leucine zippers of *fos*, *jun*, and GCN4 dictate dimerization specificity and thereby control DNA binding. Nature (London) **340**:568–571.
25. **Landschulz, W. H., P. F. Johnson, and S. L. McKnight.** 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science **240**:1759–1764.
26. **Lehming, N., J. Sartorius, M. Niemoller, G. Genenger, B. vonWilcken-Bergmann, and B. Muller-Hill.** 1987. The interaction of the recognition helix of *lac* repressor with *lac* operator. EMBO J. **6**:3145–3153.
27. **Lin, Y.-S., and M. R. Green.** 1989. Identification and purification of a *Saccharomyces cerevisiae* protein with the DNA binding specificity of mammalian activating transcription factor. Proc. Natl. Acad. Sci. USA **86**:109–113.
28. **Oliphant, A. R., C. J. Brandl, and K. Struhl.** 1989. Defining sequence specificity of DNA-binding proteins by selecting binding sites from random-sequence oligonucleotides: analysis of the yeast GCN4 protein. Mol. Cell. Biol. **9**:2944–2949.
29. **Oliphant, A. R., A. L. Nussbaum, and K. Struhl.** 1986. Cloning of random-sequence oligodeoxynucleotides. Gene **44**:177–183.
30. **O'Shea, E. K., R. Rutkowski, and P. S. Kim.** 1989. Evidence that the leucine zipper is a coiled coil. Science **243**:538–542.
31. **Passmore, S., R. Elble, and B.-K. Tye.** 1989. A protein involved in minichromosome maintenance in yeast binds a transcriptional enhancer conserved in eukaryotes. Genes Dev. **3**:921–935.
32. **Roesler, W. J., G. R. Vandenbark, and R. W. Hanson.** 1988. Cyclic AMP and the induction of eukaryotic gene transcription. J. Biol. Chem. **263**:9063–9066.
33. **Sadler, J. R., H. Sasmor, and J. L. Betz.** 1983. A perfectly symmetric *lac* operator binds the *lac* repressor very tightly. Proc. Natl. Acad. Sci. USA **80**:6785–6789.
34. **Sauer, R. T., D. L. Smith, and A. D. Johnson.** 1988. Flexibility of the yeast α2 repressor enables it to occupy the ends of its operator, leaving the center free. Genes Dev. **2**:807–816.
35. **Sellers, J. W., and K. Struhl.** 1989. Changing fos oncoprotein to a DNA-binding protein with GCN4 dimerization specificity by swapping "leucine zippers." Nature (London) **341**:74–76.
36. **Shore, D., and K. Nasmyth.** 1987. Purification and cloning of a DNA binding protein from yeast that binds to both silencer and activator elements. Cell **51**:721–732.
37. **Simons, A., D. Tils, B. vonWilcken-Bergmann, and B. Muller-Hill.** 1984. Possible ideal *lac* operator: *Escherichia coli lac* operator-like sequences from eukaryotic genomes lack the cen-

tral C:G pair. Proc. Natl. Acad. Sci. USA **81**:1624–1628.
38. **Struhl, K.** 1985. Naturally occuring poly(dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. Proc. Natl. Acad. Sci. USA **82**:8419–8423.
39. **Struhl, K.** 1985. Nucleotide sequence and transcriptional mapping of the yeast *pet56-his3-ded1* gene region. Nucleic Acids Res. **13**:8587–8601.
40. **Struhl, K.** 1986. Constitutive and inducible *Saccharomyces cerevisiae* promoters: evidence for two distinct molecular mechanisms. Mol. Cell. Biol. **6**:3847–3853.
41. **Struhl, K.** 1987. The DNA-binding domains of the jun oncoprotein and the yeast GCN4 transcriptional activator are functionally homologous. Cell **50**:841–846.
42. **Struhl, K., and D. E. Hill.** 1987. Two related regulatory sequences are required for maximal induction of *Saccharomyces cerevisiae his3* transcription. Mol. Cell. Biol. **7**:104–110.
43. **Thireos, G., M. D. Penn, and H. Greer.** 1984. 5′ untranslated sequences are required for the translational control of a yeast regulatory gene. Proc. Natl. Acad. Sci. USA **81**:5096–5100.
44. **Umesono, K., and R. M. Evans.** 1989. Determinants of target gene specificity for steroid/thyroid hormone receptors. Cell **57**:1139–1146.
45. **Vershon, A. K., R. D. Kelley, and R. T. Sauer.** 1989. Sequence-specific binding of Arc repressor to DNA: effects of operator mutations and modifications. J. Biol. Chem. **264**:3267–3273.
46. **Vogt, P. K., T. J. Bos, and R. F. Doolittle.** 1987. Homology between the DNA-binding domain of the *GCN4* regulatory protein of yeast and the carboxy-terminal region of a protein coded for by the *onc* gene *jun*. Proc. Natl. Acad. Sci. USA **84**:3316–3319.
47. **West, R. W., R. R. Yocum, and M. Ptashne.** 1984. Yeast GAL1-GAL10 divergent promoter region: location and function of the upstream activating sequence-UAS_G. Mol. Cell. Biol. **4**:2467–2478.
48. **Wobbe, C. R., and K. Struhl.** 1990. Yeast and human TATA-binding proteins have nearly identical DNA sequences requirements for transcription in vitro. Mol. Cell. Biol. **10**:3859–3867.