MOLECULAR AND CELLULAR BIOLOGY, Dec. 1989, p. 5298–5304
0270-7306/89/125298-07$02.00/0
Copyright © 1989, American Society for Microbiology
Vol. 9, No. 12

Functional Distinctions between Yeast TATA Elements

PEHR A. B. HARBURY AND KEVIN STRUHL*

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

Received 15 May 1989/Accepted 21 August 1989

Although the yeast *his3* promoter region contains two functional TATA elements, T_R and T_C, the GCN4 and GAL4 upstream activator proteins stimulate transcription only through T_R. In combination with GAL4, an oligonucleotide containing the sequence TATAAA is fully sufficient for T_R function, whereas almost all single-base-pair substitutions of this sequence abolish the ability of this element to activate transcription. Further analysis of these and other mutations of the T_R element led to the following conclusions. First, sequences downstream of the TATAAA sequence are important for T_R function. Second, a double mutant, TATTTA, can serve as a T_R element even though the corresponding single mutation, TATTAA, is unable to do so. Third, three mutations have the novel property of being able to activate transcription in combination with GCN4 but not with GAL4; this finding suggests that activation by GCN4 and by GAL4 may not occur by identical mechanisms. From these observations, we address the question of whether there is a single TATA-binding factor required for the transcription of all genes.

Most eucaryotic promoters contain a conserved sequence, TATAAA, that is located near the mRNA initiation site and is required for transcription in vivo and in vitro (13, 19a, 38). These TATA elements are specific binding sites for a protein, TFIID, required for accurate transcriptional initiation in vitro (10, 25, 31). It is generally believed that TFIID is a general factor that is part of the basic RNA polymerase II transcription machinery.

However, several observations are inconsistent with the view that TFIID is universally required for accurate transcriptional initiation by RNA polymerase II. First, approximately 20% of eucaryotic promoters lack sequences that resemble the classical TATAAA motif. Second, functionally distinct TATA elements, T_R and T_C, have been identified in the *his3* promoter by their primary sequences, interactions with upstream promoter elements, selectivity of initiation sites, and chromatin structure (7, 36). Third, deletion of the *his4* TATA element greatly reduces transcriptional activation by GCN4 but not the basal level of expression that depends on BAS1 and BAS2 (2). Fourth, overproduction of GAL4 derivatives with the acidic activation region squelches transcription dependent on the *his3* T_R but not the T_C element (12). Fifth, only some TATA sequences function in promoters activated by adenovirus E1A protein (32) or during the late stage of herpesvirus infection (16). Sixth, GCN4, normally an upstream activator protein, can stimulate transcription when its binding site replaces the normal TATA element (8).

Previously, we investigated the sequence requirements of the *his3* T_R element by saturation mutagenesis (7). In the context of a *gal-his3* hybrid promoter in which prospective T_R elements were placed downstream of an enhancer responding to GAL4 protein, an oligonucleotide containing the sequence TATAAA was sufficient for T_R function. However, 17 of the 18 possible single mutants of TATAAA abolished the ability to activate transcription in combination with GAL4, the sole exception being TATATA. We suggested that the high sequence specificity of the T_R element reflected the binding of a specific TATA-binding protein. Moreover, as the *his3* T_C element lacks a sequence that fits the T_R rules, we suggested that yeast cells might contain multiple proteins that carry out a related TATA function. By analogy, procaryotes contain multiple σ factors that interact with core RNA polymerase to generate holoenzymes that recognize distinct sequences (14).

The previous analysis of the T_R element was limited because mutations were confined to the TATAAA sequence and were assayed for functional ability only in combination with GAL4. In this paper, we address these issues by analyzing the phenotypes of single mutations of the sequence TATATA, the only single mutation of TATAAA that functions as a T_R element, as well as point mutations downstream of the TATAAA sequence. In addition, we determine whether the series of T_R derivatives can function in combination with different upstream sequences; these include a GCN4-binding site (17) as well as a poly(dA-dT) sequence that mediates constitutive transcription of the *his3* gene (34). From these experiments, we identify an exceptional double mutant, TATTTA, that functions as a T_R element and describe three mutations that function with GCN4 but not with GAL4. These results place further constraints on the hypothesis of a single TATA-binding factor required for transcription from all promoters, and they suggest that GCN4 and GAL4 may not activate transcription by identical mechanisms.

MATERIALS AND METHODS

New mutations of the *his3* T_R element. Single-base-pair substitutions in the T_R element were generated with degenerate oligonucleotides as described previously (7; Fig. 1). Two oligonucleotides were synthesized, both containing mutagenized versions of the T_R region flanked by *Eco*RI and *Sac*I recognition sequences. In one case, the sequence TATATAGTAA was mutagenized at a frequency of 9% per position (3% for each of the possible nucleotides); in the other case, the sequence was TATAAAXTAA (where X represents an equimolar mixture of A, C, and T). The oligonucleotides were converted to the double-stranded form by mutually primed synthesis (24), cleaved with *Eco*RI and *Sac*I, and inserted between the *GAL1,10* enhancer region (which contains four binding sites for GAL4 activator

* Corresponding author.

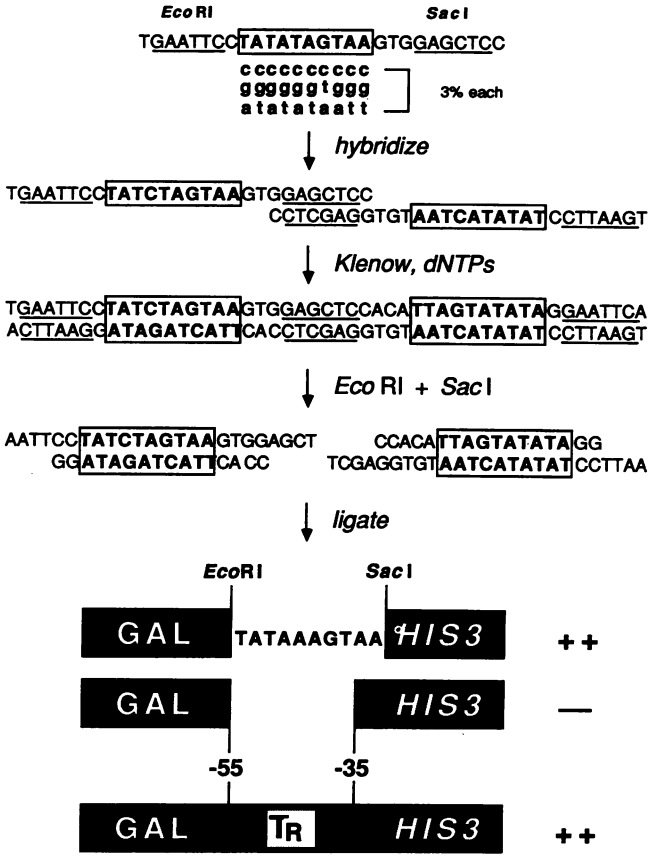

FIG. 1. Construction of mutations in the *his3* T_R element. A degenerate oligonucleotide containing the T_R sequence (boxed) flanked by *Eco*RI and *Sac*I sites at the 5′ and 3′ ends (underlined) was synthesized with 3% of each non-wild-type nucleotide (lowercase) at each of the relevant positions. This mixture of single-stranded oligonucleotides was converted to double-stranded *Eco*RI-*Sac*I fragments by the mutually primed synthesis procedure (24) and cloned between the *gal* enhancer (which contains four GAL4-binding sites) and position −24 of the *his3* gene. The resulting molecules contain the TATAAA-related sequence at the precise location in the wild-type *his3* promoter (−45 to −40) and are equivalent to previously characterized *gal-his3* promoters containing point mutations in the T_R element (7). Also indicated are the structures of related *gal-his3* promoters in which the *gal* enhancer is fused to position −55 or −35 (containing or lacking, respectively, the entire T_R element).

protein) and position −24 of the *his3* gene. The resulting DNAs were introduced into *Escherichia coli*, and those with single-base-pair substitutions were identified by DNA sequencing. All of these *G17* derivatives contain the TATAAA-related sequence at the precise location in the wild-type *his3* promoter (−45 to −40) and are equivalent to previously characterized *gal-his3* promoters containing point mutations in the T_R element (7).

Combining T_R mutations with *his3* upstream promoter elements. Various T_R derivatives were fused to position −83 or −109 of the *his3* promoter via the *Eco*RI site. Specifically, *Eco*RI-*Xho*I fragments containing the T_R allele and *his3* structural gene were inserted in place of the analogous fragments of YIp55-Sc2884 and YIp55-Sc2888 (33) to generate *his3-Δ93* and *his3-Δ94* derivatives. Both sets of derivatives delete the T_C region but retain the poly(dA-dT) element and all sequences further upstream. The *Δ93* derivatives retain the GCN4-binding site, whereas the *Δ94* derivatives do not (17).

Phenotypic assays. DNAs were introduced into strains KY320 (relevant genotype, *ura3-52 his3-Δ200 GAL*$^+$) (7) or KY329 (relevant genotype, *ura3-52 his3-TRP1 gcn4-Δ1*) (35) by replacing the *his3* locus. Growth of the resulting strains was tested in the presence of aminotriazole (AT), a competitive inhibitor of the *his3* gene product. Although the absolute level of AT resistance depends on the upstream promoter elements, the data in Tables 1 to 3 are presented so that for any set of derivatives (*G17*, *his3-Δ93*, or *his3-Δ94*), the phenotype conferred by the wild-type T_R oligonucleotide (TATAAAGTAA) was defined as ++. For activation by GAL4 (*G17* alleles; Tables 1 and 2), cells were grown on galactose, and the phenotypes were defined as follows: +++, normal growth in 40 mM AT; ++, normal growth in 20 mM AT; +, slow growth in 20 mM AT but normal growth in 10 mM AT; ±, slow growth in 5 and 10 mM AT; and −, no growth even in 5 mM AT. For activation by GCN4 (*Δ93* alleles; Table 3), cells were grown on glucose, and the phenotypes were defined as follows: ++, slow growth in 40 mM AT but normal growth in 20 mM AT; +, slow growth in 20 mM AT but normal growth in 10 mM AT; ±, slow growth in 5 and 10 mM AT; and −, no growth even in 5 mM AT. The phenotypes of these *Δ93* derivatives were unchanged when strains were grown in galactose. For constitutive transcription in the absence of either GCN4 protein (*Δ93* alleles) or the GCN4-binding site (*Δ94* alleles; Table 3), the AT concentration was reduced to 2 mM to reflect the lower expression levels. As shown previously, AT resistance is directly related to *his3* mRNA levels (7, 18, 33, 35).

RNA analysis. To measure activation by GAL4, KY320 strains with various *his3-G17* derivatives were grown in broth containing 2% galactose. To measure activation by GCN4 or by the constitutive *his3* upstream element(s), KY329 *his3-Δ93* strains that either did or did not harbor YCp88-GCN4, a plasmid capable of constitutive expression of GCN4 (18), were grown in yeast extract-peptone broth containing 2% glucose. In all cases, total 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 separated by electrophoresis in denaturing gels (7, 8). The amount of *his3* RNA was quantitated with respect to the internal *ded1* control by scanning appropriately exposed autoradiograms with a densitometer.

RESULTS

Most mutations of the sequence TATATA eliminate T_R function. In a *gal-his3* hybrid promoter whose function depends on the *his3* T_R element, an oligonucleotide containing the sequence TATAAA activates transcription in combination with GAL4 (7; Fig. 1). Of the 18 single mutations of TATAAA, only one, TATATA, conferred T_R function. Not surprisingly, 9 of 10 double mutations of TATAAA tested also failed to activate transcription, the sole exception being TATCTA. One explanation for the unexpected phenotype conferred by TATCTA is that the optimal sequence for a TATA element is actually TATATA. In this view, the exceptional double mutant would actually be interpreted as a single mutant, and the nonfunctional single mutants would be interpreted as double mutants. Although several observations argued against this possibility (7), we tested it more explicitly by analyzing single mutations from the sequence TATATA.

A degenerate oligonucleotide in which the bases TATAT AGTAA were mutated at a frequency of 9% per position was

TABLE 1. Sequences and phenotypes of mutations in the hexanucleotide core

Relevant sequence[a]	Allele	DNA fragment	Phenotype[b]	
			Glucose	Galactose
No TATA	*his3-G3*	Sc3304	–	–
TATAAA	*his3-G17*	Sc3640	–	++
TATATA	*his3-G17,215*	Sc3641	–	+
AATATA	*his3-G17,229*	Sc3720	–	–
GATATA	*his3-G17,230*	Sc3721	–	–
TCTATA	*his3-G17,231*	Sc3722	–	–
TGTATA	*his3-G17,232*	Sc3723	–	–
TTTATA	*his3-G17,233*	Sc3724	–	–
TATCTA	*his3-G17,220*	Sc3652	–	–
TATGTA	*his3-G17,228*	Sc3672	–	–
TATTTA	*his3-G17,234*	Sc3725	–	++
TATACA	*his3-G17,213*	Sc3642	–	–
TATAGA	*his3-G17,214*	Sc3643	–	–
TATATC	*his3-G17,235*	Sc3726	–	–
TATATT	*his3-G17,236*	Sc3727	–	–

[a] Mutated bases are underlined.
[b] Of strains grown in glucose or galactose minimal medium (see Materials and Methods).

cloned between the *gal* enhancer and the *his3* structural gene as described previously (7; Fig. 1). DNAs representing 12 single mutations of TATATA were introduced into yeast cells by gene replacement and analyzed for the level of *his3* transcription by growth in the presence of AT (Table 1) and by direct measurements of RNA (Fig. 2). By these analyses, 11 of the TATATA mutations were unable to activate transcription in combination with GAL4. Unexpectedly, the TATCTA derivative was inactive (we now believe the original characterization to be incorrect because of a mixup of DNAs). However, the double mutant TATTTA did function as a T_R element because it permitted growth in the presence of AT in medium with galactose but not glucose. Quantitative analysis indicated that TATTTA activated transcription as well as did TATAAA or about twice as well as did TATATA. The observation that almost all single mutations of TATATA were nonfunctional indicates that this sequence does not represent a better consensus than TATAAA and further that the TATTTA derivative is exceptional. In this regard, it should be noted that the corresponding single mutant TATTAA is nonfunctional (7).

Mutations downstream of the hexanucleotide core affect T_R function. In the wild-type *his3* T_R element, the nucleotides GTAA directly follow the TATAAA core. To investigate the role of these residues, mutations were examined for their effects on transcription as described above (Table 2; Fig. 2). Alteration of the G just beyond the core to an A or T (*his3-237* and *his3-239*) actually increased the level of *his3* transcription about two- to threefold, whereas mutation to a C (*his3-238*) essentially abolished function (only 2% of wild-type activity). These mutations conferred similar phenotypes in derivatives with a TATATA core. Thus, the position immediately downstream of the core strongly affects transcription, with A and T residues being optimal, G being tolerated, and C being severely defective.

In the context of the TATATA core, mutations at the more downstream positions could also affect T_R function. In general, mutations to C or G reduced transcription, although those at the furthest downstream position appeared to confer weaker effects. A mutation from A to T at the third position downstream did not alter the level of transcription. These results indicate that the sequence TATAAA is not sufficient

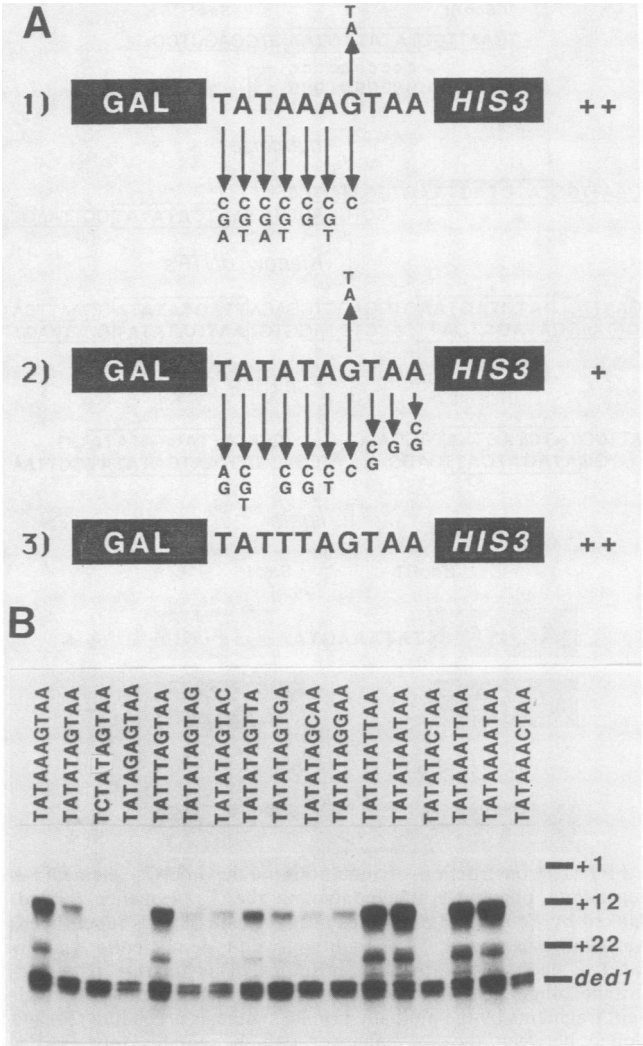

FIG. 2. (A) Phenotypic analysis of strains containing *his3-G17* derivatives. Derivatives contain the indicated sequence between the *gal* enhancer and the *his3* structural gene (see Fig. 1) and confer phenotypes described more fully in Tables 1 and 2. Symbols: ↓, mutations that decrease function (the length of the arrow reflects the severity); ↑, mutations that increase function. (1) Single-base-pair substitutions within or downstream of the TATAAA core; mutations at the three downstream most positions were not analyzed. (2) Single-base-pair substitutions within or downstream of the TATATA core; not all possible mutations were analyzed. (3) The unusual double mutation TATTTA. (B) Autoradiogram in which RNA levels were examined in galactose-grown strains containing the indicated T_R alleles. Positions of *his3* RNAs initiated at positions +1, +12, and +22 and the *ded1* internal control are indicated.

to confer T_R function and that sequences downstream of this core can play an important role. However, the downstream nucleotides do not appear to be as important as the core nucleotides because some base pairs are tolerated and because the magnitude of the defects is less.

Function of the TATA derivatives in combination with GCN4. DNAs containing all 18 of the single mutations of TATAAA (7) as well as the TATATAC allele were fused downstream of the GCN4-binding site in the *his3* promoter (17; Fig. 3). In the resulting *his3*-(*Δ93* derivatives, T_C is deleted and the distance between GCN4 and T_R (normally 53 base pairs [bp]) is reduced to 23 bp. The DNAs were

TABLE 2. Sequences and phenotypes of mutations downstream of the hexanucleotide core

Relevant sequence[a]	Allele	DNA fragment	Phenotype[b]	
			Glucose	Galactose
TATATAGTAA	*his3-G17,215*	Sc3641	−	+
TATATAATAA	*his3-G17,215,237*	Sc3728	−	+++
TATATACTAA	*his3-G17,215,238*	Sc3729	−	−
TATATATTAA	*his3-G17,215,239*	Sc3730	−	+++
TATATAGCAA	*his3-G17,215,240*	Sc3731	−	±
TATATAGGAA	*his3-G17,215,241*	Sc3732	−	±
TATATAGTGA	*his3-G17,215,242*	Sc3733	−	±
TATATAGTTA	*his3-G17,215,243*	Sc3734	−	+
TATATAGTAC	*his3-G17,215,244*	Sc3735	−	+
TATATAGTAG	*his3-G17,215,245*	Sc3736	−	±
TATAAAGTAA	*his3-G17*	Sc3640	−	++
TATAAAATAA	*his3-G17,237*	Sc3737	−	+++
TATAAACTAA	*his3-G17,238*	Sc3738	−	−
TATAAATTAA	*his3-G17,239*	Sc3739	−	+++

[a] Mutated bases are underlined.
[b] Of strains grown in glucose or galactose minimal medium (see Materials and Methods).

introduced into yeast cells by gene replacement, and the resulting strains were tested for *his3* expression by the ability to grow in AT. Under these conditions, the cells contain high levels of GCN4 (15, 38).

As expected, derivatives containing a functional T_R element as defined by GAL4 activation also conferred efficient *his3* expression in combination with GCN4 (Table 3). In addition, almost all derivatives that failed to function with GAL4 were similarly defective when located downstream of a GCN4-binding site. Surprisingly, however, *his3-206* (TTTAAA), *his3-216* (TATAAG), and *his3-215,238* (TATATAC), which do not function with GAL4 (7; Fig. 2), supported *his3* expression that was only somewhat less efficient than that observed for the TATAAA or TATATA allele. Three other alleles (*his3-202, his3-208,* and *his3-212*) showed a slight increase in *his3* expression in comparison with the majority of T_R derivatives.

To demonstrate GCN4 activation directly, *his3* RNA levels were measured under conditions in which *his3* expression is gratuitous for cell growth by introducing a plasmid that expresses GCN4 under all conditions (18) into a variety of the aforementioned strains (Fig. 3). As expected, GCN4 activated transcription from the *his3* + 12 initiation site in strains containing the TATAAG or TATATAC allele to a level approximately 30% heat of strains containing TATAAA or TATATA. RNA levels were much lower in strains containing other TATA alleles tested, although it appeared that GCN4 might stimulate transcription from *his3-Δ93* derivatives containing TATACA and possibly TATAGA and TATAAT.

Effect of TATA mutations on constitutive *his3* expression. The TATA alleles were fused to the *his3* promoter at position −109 (Fig. 3). The resulting *his3-Δ94* derivatives contain the poly(dAdT) element (34) and all other upstream sequences and are similar to the *his3-Δ93* derivatives except

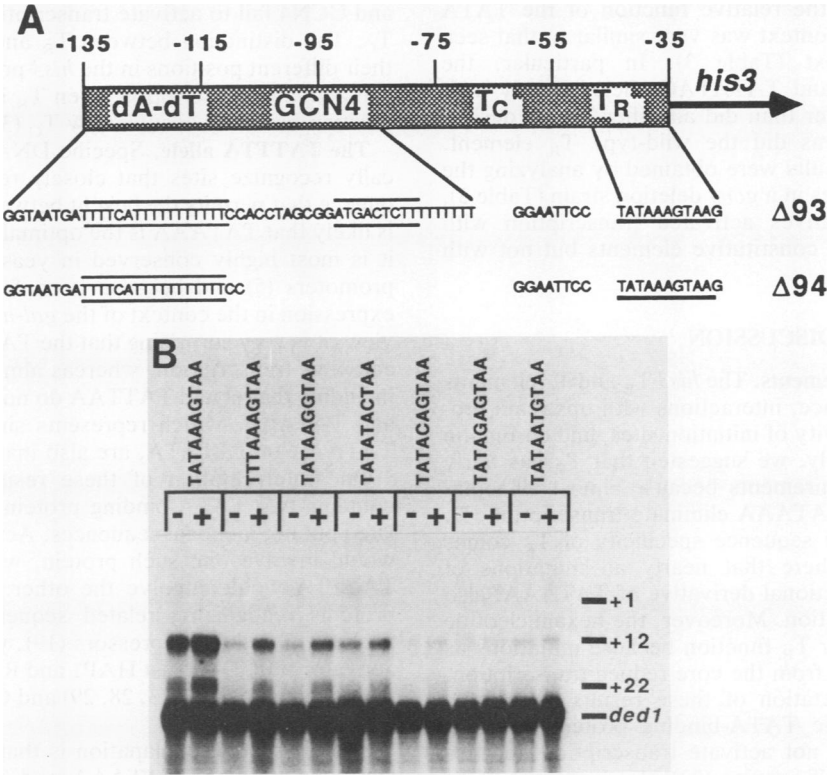

FIG. 3. Structures and RNA levels of *his3-Δ93* and *his3-Δ94* derivatives. Diagram of the *his3* promoter, with the positions of the poly(dA-dT) sequence, GCN4-binding site, and T_R and T_C TATA elements indicated with respect to the *his3* +1 mRNA initiation site. Shown below are the DNA sequences of the Δ93 and Δ94 derivatives containing the wild-type T_R oligonucleotide; an 8-bp linker containing an *Eco*RI site (GGAATTCC) lies at the junction between the *his3* upstream sequences and the T_R allele. (B) Analysis of RNAs from *gcn4-Δ1* strains containing the indicated T_R alleles that did (+) or did not (−) contain a plasmid capable of constitutive expression of GCN4 protein. Positions of *his3* RNAs initiated at positions +1, +12, and +22 and the *ded1* internal control are indicated.

TABLE 3. Effect of TATA mutations on constitutive and GCN4-activated *his3* expression

Relevant sequence[a]	*his3* allele	DNA fragment		Phenotype			
				Δ93[b]		*GCN4* (Δ94)[b]	*GAL4*[c]
		Δ93	Δ94	*gcn4-Δ1*	*GCN4*		
TATAAA	Wild type	Sc3740	Sc3760	(++)	++	(++)	++
AATAAA	*201*	Sc3741	Sc3761	−	−	−	−
CATAAA	*202*	Sc3742	Sc3762	−	±	(+)	−
GATAAA	*203*	Sc3743	Sc3763	−	−	−	−
TCTAAA	*204*	Sc3744	Sc3764	−	−	−	−
TGTAAA	*205*	Sc3745	Sc3765	−	−	−	−
TTTAAA	*206*	Sc3746	Sc3766	−	+	(+)	−
TAAAAA	*207*	Sc3747	Sc3767	−	−	−	−
TACAAA	*208*	Sc3748	Sc3768	−	±	(+)	−
TAGAAA	*209*	Sc3749	Sc3769	−	−	−	−
TATCAA	*210*	Sc3750	Sc3770	−	−	−	−
TATGAA	*211*	Sc3751	Sc3771	−	−	−	−
TATTAA	*212*	Sc3752	Sc3772	−	±	(+)	−
TATACA	*213*	Sc3753	Sc3773	−	−	−	−
TATAGA	*214*	Sc3754	Sc3774	−	−	−	−
TATATA	*215*	Sc3755	Sc3775	(+)	++	(++)	+
TATAAC	*216*	Sc3756	Sc3776	−	−	−	−
TATAAG	*216*	Sc3757	Sc3777	(±)	+	(+)	−
TATAAT	*218*	Sc3758	Sc3778	−	−	−	−
TATATAC	*215,238*	Sc3759	Sc3779	(±)	+	(+)	−

[a] Mutated bases are underlined.
[b] Strains grown in glucose medium (see Materials and Methods). Parentheses indicate growth in medium containing 2 mM AT.
[c] *G17* derivatives grown in galactose medium, phenotypes which were determined previously (7).

for a 26-bp deletion that removes the GCN4-binding site. Although deletion of the GCN4-binding site significantly decreased expression, the relative function of the TATA alleles in the *his3-Δ94* context was very similar to that seen in the *his3Δ93* context (Table 3). In particular, the TATAAG, TTTAAA, and TATATAC alleles clearly activated transcription better than did all other TATA derivatives but not as well as did the wild-type T_R element. Qualitatively similar results were obtained by analyzing the set of *his3Δ93* derivatives in a *gcn4* deletion strain (Table 3). Thus, three T_R derivatives activated transcription with GCN4 or with the *his3* constitutive elements but not with GAL4.

DISCUSSION

The *his3* T_R and T_C elements. The *his3* T_R and T_C elements differ in primary sequence, interactions with upstream promoter elements, selectivity of initiation sites, and chromatin structure (36). Previously, we suggested that T_R has fairly stringent sequence requirements because almost all mutations of the sequence TATAAA eliminate transcription (7). Further support for the sequence specificity of T_R comes from the observation here that nearly all mutations of TATATA, the sole functional derivative of TATAAA, also fail to activate transcription. Moreover, the hexanucleotide core does not suffice for T_R function because mutations as far as 4 bp downstream from the core reduce transcription.

The simplest interpretation of these results is that T_R interacts with a specific TATA-binding protein and that mutated derivatives do not activate transcription because they fail to bind. In the T_C region (33), the sequences most closely resembling T_R, CATAAT and AATGAA, differ at two positions from any of the core sequences that function with GAL4 or GCN4. If the protein interacting with T_R has strong sequence preferences, it seems unlikely that it would bind either of the TATA-like sequences in the T_C region. Moreover, if this protein could interact with these TATA-like sequences, it would be difficult to explain why GAL4 and GCN4 fail to activate transcription in combination with T_C. The distinction between T_R and T_C cannot be due to their different positions in the *his3* promoter, because GCN4 can activate transcription when T_R is moved upstream to a position normally occupied by T_C (36).

The TATTTA allele. Specific DNA-binding proteins typically recognize sites that closely resemble an optimal sequence that permits the best fit between protein and DNA. It is likely that TATAAA is the optimal core sequence because it is most highly conserved in yeast and other eucaryotic promoters (5) and because it confers the highest level of expression in the context of the *gal-his3* promoter (7). In this view, it is very surprising that the TATTTA double mutation activates transcription, whereas almost all single mutations including the related TATTAA do not. Moreover, TATAGA and TATACA, which represents single mutation of either TATAAA or TATATA, are also inactive (7).

One interpretation of these results is that yeast cells contains two TATA-binding proteins that interact with related but not identical sequences. Activation from TATAAA would involve one such protein, whereas activation from TATTTA would involve the other. Examples of multiple proteins recognizing related sequences include bacteriophage λ *cI* and *cro* repressors (19), yeast GCN4 and yAP-1 activators (18,23), yeast HAP1 and RC2 proteins (1, 26), and the mammalian AP-1 (3, 28, 29) and CCAAT (11, 20) protein families.

The alternative explanation is that a single protein interacts with both the TATAAA and TATTTA alleles. As a TATA-binding protein is unlikely to have absolute sequence specificity, TATTTA might represent an unusual configuration that is tolerated. The TATTTA sequence could have some structural similarity to TATAAA, or specific mutations within the element might have compensating effects on

function. Another possibility is that a single factor recognizes dissimilar sequences, as proposed for eucaryotic transcriptional activator proteins such as HAP1 (21), glucocorticoid receptor (30), CEBP (21), and TEF-1 (9).

Derivatives that activate in combination with GCN4 but not GAL4. GCN4 and GAL4 are usually assumed to be functionally analogous activator proteins because they stimulate transcription through acidic activation regions (18, 22). However, this view is challenged by the existence of three anomalous TATA derivatives that activate transcription in combination with GCN4 but not GAL4. This effect is not due to the fact that the closest GAL4-binding site in the *G17* derivatives is about 50 bp further upstream of the T_R element than the GCN4-binding site in *his3*-Δ*93* derivatives. Reducing the distance between the GAL4 sites and T_R in the three anomalous derivatives does not lead to transcriptional activation (C. J. Brandl and K. Struhl, unpublished data).

It seems unlikely that these three TATA derivatives simply represent weak binding sites that are below the threshold for activation by GAL4 but not GCN4 because GAL4 is the stronger activator protein. LexA-GAL4 stimulates transcription more strongly than does LexA-GCN4 when bound to a LexA operator (6, 18), and *his3* RNA levels conferred by TATAAA or TATATA are fivefold higher when activated by GAL4 as compared with GCN4 (Fig. 2 and 3). More explicitly, the TATATAC allele is 30% as efficient as the wild-type allele in combination with GCN4 but only 2% as efficient in combination with GAL4. Despite GAL4 being a stronger activator than GCN4, strains containing the anomalous derivatives grow better than their *G17* counterparts as identical AT concentrations, confirming that they cause higher levels of *his3* expression.

The apparent distinction between activation by GCN4 and GAL4 could be explained by invoking two distinct TATA-binding proteins. In this view, GCN4 activation could occur with either protein, whereas GAL4 activation could occur only with one of the proteins. Differences in the recognition properties of the two TATA-binding proteins would account for why certain TATA sequences would respond to GCN4 but not GAL4. An implication of this model is that GCN4 and GAL4 are qualitatively different with respect to the ability to interact (directly or indirectly) with these distinct TATA-binding proteins.

Alternatively, activation might involve a single TATA factor that is affected differently by GCN4 and GAL4. Several specific versions of this model could be imagined. First, binding of the TATA factor to certain DNA sequences might allosterically affect the protein such that it was receptive to GCN4 but not GAL4. Second, GCN4 and GAL4 might cause different allosteric changes in the TATA factor that would affect its ability to bind certain target sequences. Third, the TATA factor might have to induce a structural change in the DNA such as melting or unwinding to stimulate transcription. GCN4 and GAL4 could differentially affect the ability of the TATA protein to carry out this function on particular DNA sequences. By any of these models, GCN4 would be a more promiscuous but less powerful activator protein than GAL4. The functional distinctions between GCN4 and GAL4 could reflect differences in their acidic activation regions or could be related to the ability of GCN4 to interact with RNA polymerase II in vitro (4).

Are there multiple proteins performing the TATA function? The results here put further constraints on the hypothesis of a single TATA-binding factor required for transcription of all genes. A universal TATA factor would have very atypical DNA-binding properties; it would recognize the T_C element and TATTTA but not almost all single mutations of the TATAAA consensus. Moreover, this universal TATA factor would have the very unusual property that its ability to act in combination with upstream activator proteins would depend on the sequence to which it was bound. Specifically, the TATA factor would not function with GCN4 and GAL4 when bound to T_C, and it would act with GCN4 but not GAL4 when bound to the anomolous T_R derivatives. Although specific explanations can be invoked for individual observations, we feel that the DNA sequence and functional distinctions between TATA elements are more easily explained by the existence of multiple proteins performing a common TATA function.

The requirement for TFIID for transcription in vitro does not argue for or against the existence of multiple TATA factors because such experiments have generally used promoters with a consensus TATA element. In addition, as the best TFIID preparations are heterogeneous, they may contain distinct proteins that perform a related function. Although biochemical evidence for multiple TATA factors is lacking, the TATA derivatives described here should be useful substrates for their identification and characterization.

ACKNOWLEDGMENTS

We thank Wei Chen for DNAs, yeast strains, and advice during the early stages of the experiments.

This work was supported by Public Health Service grant GM 30186 to K.S. from the National Institutes of Health and by a grant from the Lucille Markey trust.

LITERATURE CITED

1. **Arcangioli, B., and B. Lescure.** 1985. Identification of proteins involved in the regulation of yeast iso-1-cytochrome c expression by oxygen. EMBO J. **4**:2627–2633.
2. **Arndt, K. T., C. Styles and G. R. Fink.** 1987. Multiple global regulators control *HIS4* transcription in yeast. Science **237**: 874–880.
3. **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.
4. **Brandl, C. J., and K. Struhl.** 1989. Yeast GCN4 transcriptional activator protein interacts with RNA polymerase II *in vitro*. Proc. Natl. Acad. Sci. USA **86**:2652–2656.
5. **Breathnach, R., and P. Chambon.** 1981. Organization and expression of eukaryotic split genes coding for proteins. Annu. Rev. Biochem. **50**:349–383.
6. **Brent, R., and M. Ptashne.** 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. Cell **43**:729–736.
7. **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.
8. **Chen, W., and K. Struhl.** 1989. Yeast upstream activator protein GCN4 can stimulate transcription when its binding site replaces the TATA element. EMBO J. **8**:261–268.
9. **Davidson, I., J. H. Xiao, R. Rosales, A. Staub, and P. Chambon.** 1988. The HeLa cell protein TEF-1 binds specifically and cooperatively to two SV40 enhancer motifs of unrelated sequence. Cell **54**:931–942.
10. **Davison, B. L., J.-M. Egly, E. R. Mulvihill, and P. Chambon.** 1983. Formation of stable preinitiation complexes between eukaryotic class B transcription factors and promoter sequences. Nature (London) **301**:680–686.
11. **Dorn, A., J. Bollekens, A. Staub, C. Benoist, and D. Mathis.** 1987. A multiplicity of CCAAT box-binding proteins. Cell

50:863–872.
12. **Gill, G., and G. Ptashne.** 1988. Negative effect of the transcriptional activator GAL4. Nature (London) **334**:721–723.
13. **Guarente, L.** 1988. UASs and enhancers: common mechanism of transcriptional activation in yeast and mammals. Cell **52**: 303–305.
14. **Helmann, J. D., and M. J. Chamberlin.** 1988. Structure and function of bacterial sigma factors. Annu. Rev. Biochem. **57**: 839–872.
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. **Homa, F. L., J. C. Glorioso, and M. Levine.** 1988. A specific 15-bp TATA box promoter element is required for expression of a herpes simplex virus type I late gene. Genes Dev. **2**:40–53.
17. **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.
18. **Hope, I. A., and K. Struhl.** 1986. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. Cell **46**:885–894.
19. **Johnson, A. D., A. R. Poteete, G. Lauer, R. T. Sauer, G. R. Ackers, and M. Ptashne.** 1981. λ repressor and cro- components of an efficient molecular switch. Nature (London) **294**:217–223.
19a. **Johnson, P. F., and S. L. McKnight.** 1989. Eukaryotic transcriptional regulatory proteins. Annu. Rev. Biochem. **58**:799–839.
20. **Jones, K. A., J. T. Kadonaga, P. J. Rosenfeld, T. J. Kelly, and R. Tjian.** 1987. A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. Cell **48**:79–89.
21. **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.
22. **Ma, J., and M. Ptashne.** 1987. Deletion analysis of GAL4 defines two transcriptional activating segments. Cell **48**:847–853.
23. **Moye-Rowley, W. S., K. D. Harshman, and C. S. Parker.** 1989. Yeast *YAP1* encodes a novel form of the jun family of transcriptional activator proteins. Genes Dev. **3**:283–292.
24. **Oliphant, A. R., A. L. Nussbaum, and K. Struhl.** 1986. Cloning of random-sequence oligodeoxynucleotides. Gene **44**:177–183.
25. **Parker, C. S., and J. Topol.** 1984. A Drosophila RNA polymerase II transcription factor contains a promoter-region specific DNA-binding activity. Cell **36**:357–369.
26. **Pfeifer, K., B. Arcangiolo, and L. Guarente.** 1987. Yeast HAP1 activator competes with the factor RC2 for binding to the upstream activator site UAS1 of the *CYC1* gene. Cell **49**:9–18.
27. **Pfeifer, K., T. Prezant, and L. Guarente.** 1987. Yeast HAP1 activator binds to two upstream activation sites of different sequence. Cell **49**:19–27.
28. **Rauscher, F. J., III, D. R. Cohen, T. Curran, T. J. Bos, P. K. Vogt, D. Bohmann, R. Tjian, and B. R. Franza.** 1988. Fos-associated protein p39 is the product of the *jun* proto-oncogene. Science **240**:1010–1016.
29. **Ryder, K., L. F. Lau, and D. Nathans.** 1988. A gene activated by growth factors is related to the oncogene v-*jun*. Proc. Natl. Acad. Sci. USA **85**:1487–1491.
30. **Sakai, D. D., S. Helms, J. Carlstedt-Duke, J.-A. Gustafsson, F. M. Rottman, and K. R. Yamamoto.** 1988. Hormone-mediated repression: a negative glucocorticoid response element from the bovine prolactin gene. Genes Dev. **2**:1144–1154.
31. **Sawadogo, M., and R. G. Roeder.** 1985. Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. Cell **43**:165–175.
32. **Simon, M. C., T. M. Fisch, B. J. Benecke, J. R. Nevins, and N. Heintz.** 1988. Definition of multiple, functionally distinct TATA elements, one of which is a target in the *hsp70* promoter for E1A regulation. Cell **52**:723–729.
33. **Struhl, K.** 1982. Regulatory sites for *his3* expression in yeast. Nature (London) **300**:284–287.
34. **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.
35. **Struhl, K.** 1987. The DNA-binding domains of the jun oncoprotein and the yeast GCN4 transcriptional activator are functionally homologous. Cell **50**:841–846.
36. **Struhl, K.** 1986. Constitutive and inducible *Saccharomyces cerevisiae* promoters: evidence for two distinct molecular mechanisms. Mol. Cell. Biol. **6**:3847–3853.
37. **Struhl, K.** 1989. Molecular mechanisms of transcriptional regulation in yeast. Annu. Rev. Biochem. **58**:1051–1077.
38. **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.