Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 2691–2695, April 1988
Genetics

Saturation mutagenesis of a yeast *his3* "TATA element": Genetic evidence for a specific TATA-binding protein

(gene regulation/eukaryotic promoters/transcription/DNA-binding protein/mRNA initiation)

WEI CHEN AND KEVIN STRUHL

Department of Biological Chemistry, Harvard Medical School, Boston, MA 02115

Communicated by Charles C. Richardson, January 7, 1988

ABSTRACT **The yeast *his3* promoter region contains two distinct classes of "TATA elements," constitutive (T_C) and regulatory (T_R), that are defined by their interactions with upstream promoter elements, selectivity of initiation sites, and chromatin structure. T_C is localized between positions −83 and −53, and T_R is localized between positions −55 and −35, regions in which there are several TATA-like sequences. In this study, we used saturation mutagenesis to examine the structural requirements of the *his3* T_R element necessary for transcriptional induction. To avoid the complications of redundant elements, the phenotypic analysis was carried out by using a *gal–his3* hybrid promoter whose function depends on a short oligonucleotide containing the prospective *his3* T_R element. In this context, an oligonucleotide containing the sequence TATAAA is sufficient for T_R function. However, 17 out of the 18 possible single-base substitutions and 9 out of 10 double mutations of this sequence abolish T_R function. This strict sequence requirement for T_R function strongly suggests that the T_R element is a target site for a sequence-specific DNA-binding protein. Further, as the region encoding T_C and promoters of certain other yeast genes do not contain a sequence that is compatible with T_R function, we suggest that yeast cells contain multiple proteins with distinct sequence specificities that carry out a related "TATA function" and that yeast promoters can be divided into classes based on their downstream promoter elements.**

Most eukaryotic promoters for genes transcribed by RNA polymerase II contain a DNA sequence called the "TATA element" that has a critical role in transcription (for review, see refs. 1–3). It is commonly assumed that TATA sequences are general promoter elements that are recognized by a common transcription factor that is part of the basic transcriptional machinery. However, there are a number of genes whose promoters seem to lack TATA sequences. Moreover, there are several examples in which deletion of the TATA sequence results in a minor effect on the level of transcription. In addition, two functionally distinct TATA elements have been defined in the yeast *his3* promoter by their interactions with upstream promoter elements, selectivity of initiation sites, and chromatin structure (4). These observations are seemingly inconsistent with the common view of the role of TATA elements.

Much of the confusion about TATA elements may reflect problems in the interpretation of the experimental data. First, the nucleotide sequence requirements for a functional TATA element are poorly understood. This makes it difficult to determine whether particular A+T-rich sequences within a particular promoter region are actually TATA elements simply by inspecting the sequence. Second, eukaryotic promoters often contain multiple copies of a particular element each of which is or could be functional in an appropriate context. Such potential redundancy makes it difficult to interpret the results of mutational studies because the elimination of one element might have a minimal functional consequence. This problem is particularly evident for yeast promoters because multiple TATA-like sequences are often present and the distance to the mRNA initiation site is rather variable. Third, as postulated for the yeast *his3* promoter, there might be various kinds of TATA elements with different DNA sequence and spacing requirements. These problems, individually and in combination, make it very difficult to interpret the results of genetic experiments in molecular terms.

Transcription of the yeast *his3* gene normally occurs at a basal level of 1–2 mRNA molecules per cell, but it can be induced 3-fold under conditions of amino acid starvation. Detailed analysis of the *his3* promoter localizes the TATA function to a 48-base-pair (bp) region between positions −83 and −35 with respect to the most upstream transcription initiation site (ref. 5, see Fig. 1). Several lines of evidence indicate that this TATA region contains two distinct elements: one involved in regulated transcription (T_R) and the other involved in constitutive transcription (T_C) (4). Small deletions within the region abolish only T_R function without affecting T_C function, whereas a deletion of the whole region destroys both T_R and T_C. Transcription dependent on T_C is initiated equally from two sites, positions +1 and +12, whereas inducible transcription dependent on T_R is initiated preferentially from the +12 site; this selectivity is determined primarily by the distance between T_R and the initiation sites. Finally, T_R but not T_C can function together with the GCN4 binding site in the natural *his3* promoter (4), the GAL4 binding site from the *gal* promoter (6), and a cryptic site that is uncovered by *ope* suppressor mutations (7). From these experiments, T_C maps between positions −53 and −83, whereas T_R is localized between positions −55 to −35.

Inspection of the DNA sequence between positions −83 and −35 reveals that there are several TATA-like sequence elements (ref. 8; Fig. 1). However, without a fine structural analysis to precisely delineate the DNA sequence requirements for a functional element, it is difficult to distinguish between functional TATA elements and A+T-rich sequences and between redundant and functionally different TATA elements. Here we use saturation mutagenesis *in vitro* and functional analysis *in vivo* to dissect the structural requirements of the *his3* T_R element that is necessary for transcriptional induction.

MATERIALS AND METHODS

The procedures for DNA manipulations (11), DNA sequencing (12), yeast transformation (8), RNA preparation (8), and S1 nuclease analysis with oligonucleotide probes (13) have been described. All DNA molecules containing the "UAS_G element" are similar to *gal–his3* fusions (6) except that a *Sac*

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

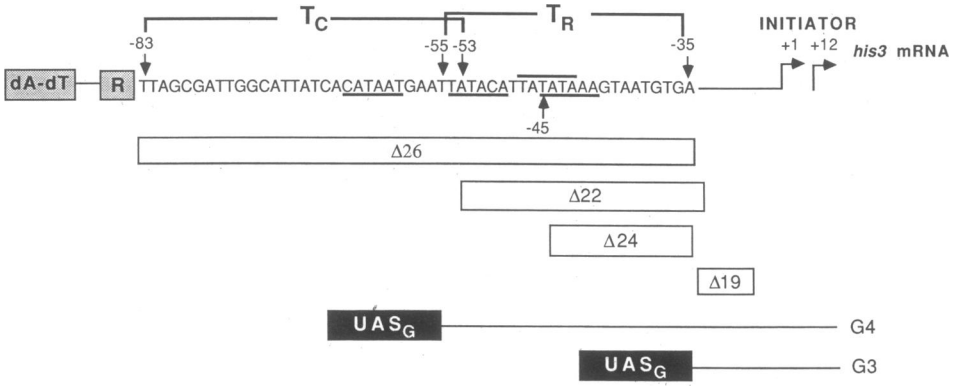

FIG. 1. *his3* promoter region of wild-type and mutant strains. The DNA sequence of the *his3* TATA region, nucleotides −35 to −83 with respect to the mRNA initiation sites at positions +1 and +12 (arrows), is shown with TATA-like sequences (under- or overlined). The shaded dA-dT box represents the upstream element for constitutive transcription (9), and the R box represents the GCN4 binding site for regulated transcription (10). T_C and T_R are the TATA elements for constitutive and regulatory *his3* transcription whose boundaries (thick brackets) have been defined by deletion mutants (open boxes) (4, 5) and *gal–his3* hybrid promoters (solid boxes) (6). *his3*–Δ26 (positions −83 to −35) removes both T_C and T_R because it abolishes both constitutive and regulatory transcription. *his3*–Δ22 (positions −53 to −34) defines the downstream border of T_C because it reduces induced transcription while maintaining the normal basal level. *his3*–Δ24 (positions −44 to −35) is the smallest deletion that destroys T_R. *his3*–Δ19, -G4, and -G3 define the location of T_R. In *his3*–Δ19, the sequence between positions −35 and −24 was replaced with an 8-bp *Eco*RI linker and transcription was unaffected. In *his3*–G4 and -G3, the 365-bp *gal* fragment was fused to positions −55 and −35, respectively; G4 is fully inducible by galactose, and G3 is not inducible.

I site was introduced at position −24. This was accomplished by ligating an 8-bp *Sac* I linker to the *Taq* I site at position −24 that had been blunted with the large fragment of DNA polymerase I. Yeast strain KY320 was constructed by replacing the *leu2* locus of KY117 (6) with the *leu2–PET56* allele. This allele was constructed by inserting the 2-kilobase *Bst*EII–*Eco*RI fragment of YRp14-Sc2850 (11) containing the entire *pet56* gene (*his3* nucleotides −35 to −2000) between the *Eco*RI and *Bst*EII sites in the *leu2* structural gene. The reason for constructing a strain with the *leu2–PET56* allele was that replacement of the *his3* locus by the *gal–his3* fusions would otherwise lead to a petite phenotype. Due to structure of the *gal–his3* fusions, the UAS_G fragment replaces most of the *pet56* promoter region and hence greatly reduces the level of *pet56* expression (9).

RESULTS

Experimental Design. The following strategy was adopted to determine the nucleotide sequence requirements for a TATA element. To eliminate the problem of redundancy, the starting molecule contained a defective promoter that could not support transcription because it lacked any functional TATA element. Then, short oligonucleotides containing a presumptive TATA element were inserted into this defective promoter to identify a minimal sequence that was sufficient for wild-type TATA function. Once identified, this minimal functional sequence was saturated with base-pair substitutions by cloning degenerate oligonucleotides, and the resulting molecules were tested for their ability to activate transcription.

We utilized the properties of the *gal–his3* hybrid promoters (6) to develop a simple assay for T_R function (Fig. 2). Specifically, oligonucleotides containing prospective T_R elements were cloned between UAS_G and the *his3* mRNA coding sequences. By choosing an appropriate fusion point, the control molecule lacking an oligonucleotide would not activate transcription due to the absence of any functional TATA element. Insertion of an oligonucleotide containing a functional T_R element should confer high levels of *his3* transcription in galactose medium but not in glucose medium. Moreover, by comparing the level of *his3* transcription and the selectivity of initiation sites conferred by such alleles to that observed in fully functional *gal–his3* fusions, T_R function conferred by any particular oligonucleotide can be measured quantitatively.

TATAAA Is a Fully Functional T_R Element. The 20-bp region within which the T_R element has been localized contains the sequence TATAAA between positions −45 and −40. This sequence, which is identical to the TATA consensus sequence, occurs only once in the *his3–pet56* promoter region, and small deletions that remove this sequence abolish T_R function. In addition, a TATAAA oligonucleotide was functional in the context of the *his4* promoter (14). Thus, we decided to test whether TATAAA was sufficient to encode the *his3* T_R element.

A degenerate oligonucleotide containing TATAAA and seven nucleotides downstream from it and flanked by *Eco*RI and *Sac* I sites at the 5′ and 3′ ends was synthesized. The seven additional nucleotides are identical to those found in the native *his3* promoter downstream from the TATAAA sequence. By using mutually primed synthesis (ref. 15; Fig.

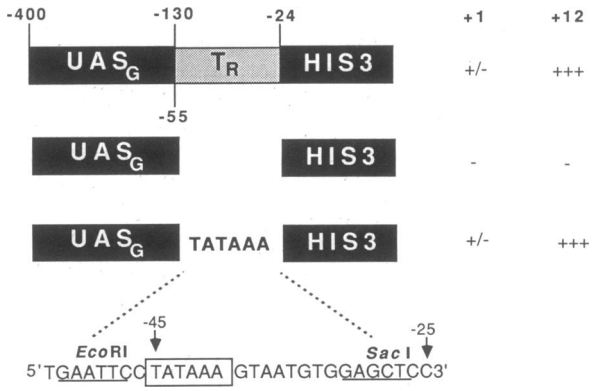

FIG. 2. Experimental design. The 365-bp *gal* fragment UAS_G is fused between positions −400 and −130 of the *his3* gene promoter region (6). (*Top*) Structure of *his3*–G4, a fusion to position −55 that contains a fully functional T_R element. (*Middle*) Starting molecule for the mutagenesis experiments with the *gal* fragment fused to position −24 (i.e., no TATA element). (*Bottom*) *Eco*RI–*Sac* I oligonucleotide with degenerate nucleotides at the position that the TATAAA (box) was inserted between the *gal* fragment and position −24 of *his3* gene. The levels of his3 RNA initiated from the +1 and +12 sites in galactose-grown strains containing these *gal–his3* fusions are indicated qualitatively to the right.

3), this oligonucleotide was inserted between UAS_G (a 365-bp fragment from the *GAL1,10* promoter that confers galactose induction) and position −24 of the *his3* gene. The position −24 insertion site was chosen so that the TATAAA sequence in our construction would be at the same distance (positions −40 to −45) from the start site of transcription as the TATAAA in the native *his3* promoter. The resulting molecule YIp55-Sc3640 was identical to previous *gal–his3* fusions except that all *his3* sequences upstream of TATAAA were eliminated and a *Sac* I linker replaced the region between positions −31 and −25.

YIp55-Sc3640 was introduced into the yeast strain KY320 such that it replaced the original chromosomal *his3* locus. The resulting strain was analyzed for the level of *his3* expression by its growth properties in the presence of aminotriazole (a competitive inhibitor of imidazole glycerol phosphate dehydratase, the *his3* gene product) and by quantitative measurements of imidazole glycerol phosphate dehydratase and his3 RNA (Fig. 4). By all of these assays, the *his3*–G17 allele containing the TATAAA oligonucleotide behaves very similarly to *his3*–G4, an allele in which UAS_G is fused to position −55 and hence contains a fully functional T_R element (6). Upon galactose induction, it confers equally efficient growth in the presence of 20 mM aminotriazole, produces equally high levels of protein product (data not shown), and makes equal levels of mRNA transcripts (compare lanes 6 and 7 of Fig. 4). Moreover, galactose induction of *his3* transcription occurs with a strong preference at the +12 and +22 sites, just as is observed in other *gal–his3* fusions (6). As expected, his3 induction does not occur in the absence of the *his3* TATAAA oligonucleotide (lane 5) nor when cells are grown in glucose medium (lanes 2–4). Thus, the results suggest that TATAAA is functionally equivalent to the *his3* T_R element.

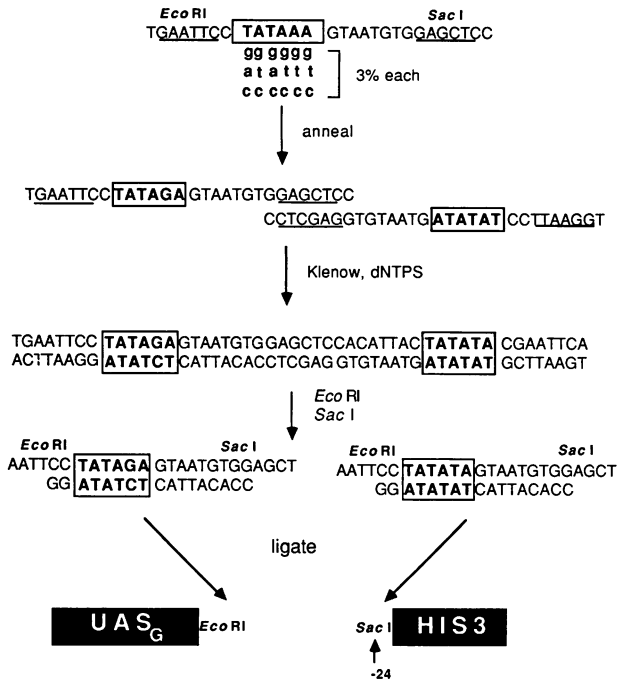

FIG. 3. Construction of TATA mutations. Degenerate oligonucleotide sequence containing *Eco*RI and *Sac* I sites at the 5′ and 3′ ends and the TATAAA sequence that was synthesized with 3% of each mutant nucleotide inserted at each position is shown at the top. Shown below is the mutually primed synthesis procedure (15) that converted this mixture of single-stranded oligonucleotides into the double-stranded form with *Eco*RI and *Sac* I sites suitable for cloning into the *gal–his3* fusion molecule shown at the bottom.

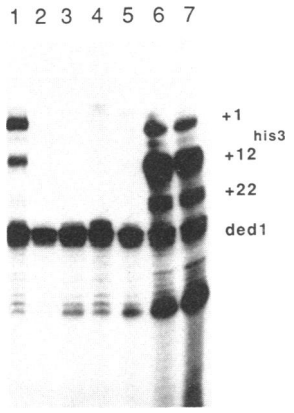

FIG. 4. Quantitation of his3 RNA levels by S1 nuclease analysis. RNAs from yeast strains grown in glucose (lanes 1–4) or galactose (lanes 5–7) medium containing the following *his3* alleles were hybridized to completion with an excess of ^{32}P-labeled oligonucleotides for *his3* and the *ded1* internal control (13): wild-type *his3* (lane 1); *his3*–G3, fusion to position −35 that lacks a TATA element (lanes 2 and 5); *his3*–G4, fusion to position −55 that contains the fully functional T_R element (lanes 3 and 6); and *his3*–G17, *gal–his3* fusion containing the TATAAA oligonucleotide (lanes 4 and 7). The positions of his3 RNAs initiated at positions +1, +12, and +22 and the ded1 control RNA are indicated.

Almost all Single-Base Substitutions of TATAAA Abolish Transcription. Having established that TATAAA is a functional T_R element, we applied a saturation mutagenesis procedure using degenerate oligonucleotides to determine the structural requirements of this sequence. The degenerate oligonucleotide described above was synthesized such that there was a 10% mutation rate at each position of the TATAAA sequence. DNA sequence analysis of 43 molecules containing an inserted oligonucleotide yielded 12 single-base substitutions and 10 double-base substitutions. To obtain the remaining six possible single mutations, two appropriately designed degenerate oligonucleotides were cloned in the same manner. All these mutations were introduced into yeast and phenotypically analyzed in the same way as the TATAAA allele described in the above section.

To our surprise, 17 out of the 18 possible single mutations of TATAAA conferred very low levels of *his3* expression as evidenced by the failure of strains to grow in aminotriazole (Table 1) and by direct RNA measurements (Fig. 5). The phenotypes were essentially similar to that observed with *his3*–G3, the derivative that lacked the oligonucleotide, although some derivatives showed a very slight increase in *his3* expression (strains grew faster in galactose medium lacking aminotriazole; Table 1). The only functional sequence out of the single substitutions was TATATA, which resulted in approximately half of the *his3* RNA level as compared to that achieved by TATAAA. Given the results of the single mutations, it was not surprising to find that 9 out of the 10 double mutations were defective for T_R function. Interestingly, however, the double mutation TATCTA behaved similarly to TATATA. For both the TATATA and TATCTA derivatives, *his3* transcription was initiated preferentially at the +12 and +22 sites in galactose grown cells, and it was not observed in glucose grown cells.

DISCUSSION

DNA Sequence Requirements for the *his3* T_R Element. Fusion of UAS_G to the TATAAA sequence at position −45 yields a promoter that behaves indistinguishably from other *gal–his3* fusions. This result indicates that sequences upstream of position −45 are unnecessary for T_R function, and it strongly suggests that TATAAA is sufficient as a fully

Table 1. Sequences and phenotypes of TATA point mutations

Sequence	Allele	Fragment	Glucose		Galactose		
			− AT	5 mM AT	− AT	10 mM AT	20 mM AT
Wild-type	*HIS3*	Sc2812	++	++	++	++	++
$UAS_G/-55$	*his3*–G4	Sc3305	+	−	++	++	++
$UAS_G/-35$	*his3*–G3	Sc3304	+	−	+	−	−
TATAAA	*his3*–G17	Sc3640	+	−	++	++	++
*A*ATAAA	*his3*–G17,201	Sc3689	+	−	+	−	−
*C*ATAAA	*his3*–G17,202	Sc3669	+	−	++	−	−
*G*ATAAA	*his3*–G17,203	Sc3688	++	−	++	−	−
T*C*TAAA	*his3*–G17,204	Sc3691	+	−	+	−	−
T*G*TAAA	*his3*–G17,205	Sc3660	±	−	+	−	−
T*T*TAAA	*his3*–G17,206	Sc3653	+	−	++	−	−
TA*A*AAA	*his3*–G17,207	Sc3656	+	−	+	−	−
TA*C*AAA	*his3*–G17,208	Sc3687	+	−	++	−	−
TA*G*AAA	*his3*–G17,209	Sc3651	+	−	++	−	−
TAT*C*AA	*his3*–G17,210	Sc3690	±	−	++	−	−
TAT*G*AA	*his3*–G17,211	Sc3666	+	−	++	−	−
TAT*T*AA	*his3*–G17,212	Sc3662	+	−	++	−	−
TATA*C*A	*his3*–G17,213	Sc3642	+	−	++	−	−
TATA*G*A	*his3*–G17,214	Sc3643	±	−	±	−	−
TATA*T*A	*his3*–G17,215	Sc3641	+	−	++	++	++
TATAA*C*	*his3*–G17,216	Sc3649	+	−	+	−	−
TATAA*G*	*his3*–G17,217	Sc3686	+	−	+	−	−
TATAA*T*	*his3*–G17,218	Sc3665	+	−	+	−	−
T*G*TAA*G*	*his3*–G17,219	Sc3644	+	−	+	−	−
TATCTA	*his3*–G17,220	Sc3652	+	−	++	++	++
TATT*AG*	*his3*–G17,221	Sc3654	+	−	+	−	−
*CG*TAAA	*his3*–G17,222	Sc3655	+	−	+	−	−
*GA*TTAA	*his3*–G17,223	Sc3663	+	−	+	−	−
*GA*TCAA	*his3*–G17,224	Sc3667	+	−	+	−	−
TA*GAAG*	*his3*–G17,225	Sc3668	+	−	+	−	−
T*TC*AAA	*his3*–G17,226	Sc3670	+	−	+	−	−
TATA*CT*	*his3*–G17,227	Sc3671	+	−	+	−	−
TAT*GTA*	*his3*–G17,228	Sc3672	+	−	+	−	−

For each derivative, the relevant DNA sequence (mutated bases in italics), *his3* allele number, and DNA fragment number are listed. To determine phenotypes, strains were grown in glucose or galactose minimal medium lacking histidine in the absence or presence of 5 mM, 10 mM, or 20 mM aminotriazole (AT). The phenotypes listed are as follows: ++, grows indistinguishably from the wild-type strain; +, grows slower than wild-type; ±, grows very slowly; −, does not grow. In medium lacking aminotriazole, strains that have his3 RNA levels ≈5% of his3 mRNA basal level will grow slowly, whereas those having >10% of the basal his3 mRNA level will grow at wild-type rates. The basis of the residual *his3* expression that permits all strains to grow in the absence of aminotriazole is unknown.

functional T_R. The region between positions −39 and −35 has not been excluded for a possible role in T_R function, although these five nucleotides are not conserved. In contrast, the sequence TATAAA is frequently observed in yeast and other eukaryotic promoter regions relatively close to the transcriptional initiation site.

The facts that 17 out of the 18 possible single-base substitutions and 9 out of 10 double mutations of the TATAAA sequence abolish transcription indicate that the sequence specificity for T_R function is extremely high. Moreover, these results demonstrate that it is possible to change a highly efficient promoter into an inactive promoter by single base-pair changes. These observations are surprising in light of the general view that many TATA-like sequences can act as functional promoter elements and that mutations or deletions of the TATA sequence often do not eliminate transcription. We feel that these apparent differences reflect the presence of redundant or other genetic

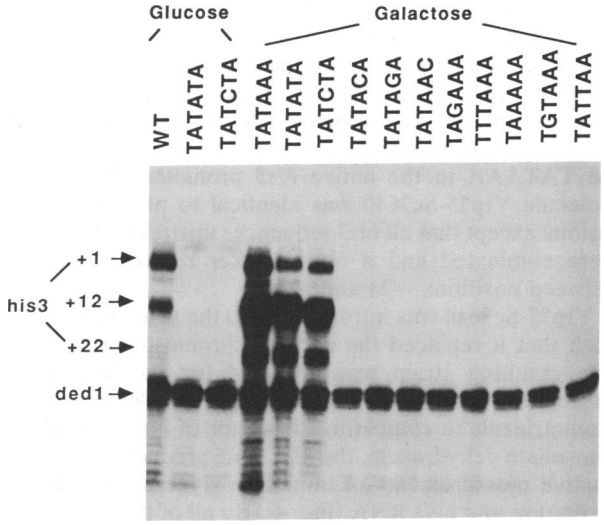

FIG. 5. Quantitation of his3 RNA levels in strains containing representative TATA point mutants. RNAs from 10 single-point mutants and from the TATAAA strain grown in galactose medium and RNAs from the TATATA and TATCTA strains grown in glucose medium were analyzed as described in Fig. 4. Lanes are labeled by the name of the oligonucleotide used.

elements in the other promoters that have been examined. By creating a situation in which all potential TATA elements are eliminated and full function is dependent on a short oligonucleotide, the high sequence specificity of the T_R element could be observed.

Although almost all mutations of TATAAA eliminate T_R function, the single mutant TATATA and the double mutant TATCTA promote *his3* transcription at levels almost as high as TATAAA. One consequence of this is that the native *his3* promoter is redundant for the T_R function because it contains the sequence TATATAAA between nucleotides −47 and −40. The phenotype conferred by the exceptional double mutation is surprising particularly because the related single mutation TATCAA shows no detectable T_R function.

One explanation is that the TATCTA double mutation fortuitously generates a sequence that is recognized by a different protein that performs a related function (see below). Less likely, this double mutation might create an unusual configuration that is tolerated by the putative T_R binding protein. Another explanation is that the consensus T_R element is actually TATATA, not TATAAA. In this view, most single mutations in this paper would actually be double mutations, and the TATCTA double mutation would actually be a single mutation. However, this interpretation does not explain why TATAAA is the more highly conserved sequence, why TATATA is less functional than TATAAA, or why the "single mutations" TATAGA, TATACA, and TATGTA are not functional. Nevertheless, this kind of explanation points out the main flaw in defining the DNA sequence requirements of a genetic element by mutagenesis of a single sequence, namely the inability to obtain related sequences that may confer a similar function. To avoid this problem and to determine the requirements more precisely, it might be useful to employ the random selection method (16).

Evidence for a Specific Protein That Binds to the T_R Element. Proteins that bind specifically to TATA elements have been identified in extracts from flies and mammals (17–19). The transcription factor IID, which is required for accurate initiation *in vitro*, binds tightly to a 10-bp region centered on the consensus TATA sequence while superficially contacting four turns of the DNA helix downstream

from the TATA box (18). In yeast cells, the only evidence for a TATA-binding activity comes from "photofootprinting" *in vivo* in which enhancement of the ultraviolet-induced, covalent modification at a thymidine within the putative *GAL1,10* TATA element was observed upon transcriptional activation (20).

The high sequence specificity of the *his3* T_R element provides strong evidence for a protein that binds to this sequence and facilitates transcription. Furthermore, we propose that this putative T_R-binding protein is indeed defined by its DNA sequence requirements. Although we have no information about the physical properties of such a protein (e.g., molecular weight or chromatographic behavior), its rigid sequence requirements should be sufficient to distinguish this protein from others. In this regard, the mutants generated in this study should be useful for identifying and characterizing the T_R protein in biochemical terms. Besides providing substrates for DNA-binding experiments, the TATA mutations might be useful for obtaining revertants that grow in aminotriazole in the hope that transcription can be restored through compensatory mutations in the putative T_R-binding protein.

It is possible that the putative protein defined by mutational analysis is the yeast analog of transcription factor IID. Transcription factor IID and an upstream stimulating factor bind simultaneously to the promoter, suggesting that direct contact between these two proteins is necessary for transcriptional initiation (19). Similarly, it has been proposed from genetic experiments that yeast activator proteins, such as GCN4 and GAL4, need to interact with T_R for transcriptional activation (2). However, the mechanism of initiation by T_R protein might differ from transcription factor IID because it appears that yeast TATA elements, unlike their higher eukaryotic counterparts, do not position mRNA start sites (2, 3).

Evidence for Multiple Proteins with TATA-Like Function. The standard conception of eukaryotic TATA elements includes two distinct aspects, a conserved sequence and a required promoter element located close to the mRNA initiation site. The proposal of a specific T_R-binding protein with a high sequence specificity has direct implications for yeast promoters that suggest that these two aspects should be separated. The majority of yeast promoters contain the sequence TATAAA or TATATA near mRNA start sites. For this class, it is extremely likely that the T_R protein will be involved in the transcriptional initiation process. However, certain yeast promoters do not contain sequences that are compatible with T_R function even though they may contain A+T-rich sequences. Assuming that these promoters require "downstream elements" for transcription, it follows that such elements cannot be recognized by the T_R protein but rather are targets for another DNA-binding protein. Transcription of the *his3* gene provides a clear example of this idea because the region containing the T_C element necessary for basal level expression does not have a sequence that fits the T_R rules. Thus, the results of this paper strongly suggest that the T_R and T_C elements, although close together and functionally related, interact with different proteins. This also provides the simplest explanation for the functional distinctions between T_R and T_C elements with regard to their interactions with upstream activator proteins.

From these considerations, we suggest that yeast cells contain multiple proteins that carry out a related "downstream element function" but have different sequence specificities for DNA binding. This is reminiscent of the situation involving bacterial σ factors that interact with the core RNA polymerase to generate distinct holoenzymes that recognize different -10 and -35 promoter sequences (21). By this analogy, we suggest that T_R is similar to the 70-kDa σ factor in that both proteins are general transcriptional factors that are important for the majority of genes in the organism. In this view, proteins that recognize T_C or other sequences would be analogous to the minor σ factors. In fact, it seems likely that various TATA elements might not bear any sequence resemblance to the sequence TATAAA at all. We emphasize that the above analogy does not reflect mechanistic similarities between yeast TATA-binding proteins and bacterial σ factors. However, the proposal that yeast (and presumably other eukaryotic) promoters can be divided into classes based on their downstream promoter elements is useful for explaining many apparently contradictory observations and for understanding how specific TATA-binding proteins perform their functions.

This work was supported by a grant from the National Institutes of Health (GM 30186).

1. McKnight, S. L. & Tjian, R. (1987) *Cell* **46**, 795–805.
2. Struhl, K. (1987) *Cell* **49**, 295–297.
3. Guarente, L. (1987) *Annu. Rev. Genet.* **21**, 425–452.
4. Struhl, K. (1986) *Mol. Cell. Biol.* **6**, 3847–3853.
5. Struhl, K. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 7385–7389.
6. Struhl, K. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 7865–7869.
7. Oettinger, M. A. & Struhl, K. (1985) *Mol. Cell. Biol.* **5**, 1901–1909.
8. Struhl, K. (1985) *Nucleic Acids Res.* **13**, 8587–8601.
9. Struhl, K. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 8419–8423.
10. Hope, I. A. & Struhl, K. (1985) *Cell* **43**, 177–188.
11. Struhl, K. (1983) *Gene* **26**, 231–241.
12. Chen, E. Y. & Seeburg, P. H. (1985) *DNA* **4**, 165–170.
13. Chen, W., Tabor, S. & Struhl, K. (1987) *Cell* **50**, 1047–1055.
14. Nagawa, F. & Fink, G. R. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 8557–8561.
15. Oliphant, A. R. & Struhl, K. (1986) *Gene* **44**, 177–183.
16. Oliphant, A. R. & Struhl, K. (1987) *Methods Enzymol.* **155**, 568–582.
17. Davison, B. L., Egly, J.-M., Mulvihill, E. R. & Chambon, P. (1983) *Nature (London)* **301**, 680–686.
18. Parker, C. S. & Topol, J. (1984) *Cell* **36**, 357–369.
19. Sawadogo, M. & Roeder, R. G. (1985) *Cell* **43**, 165–175.
20. Selleck, S. B. & Majors, J. (1987) *Nature (London)* **325**, 173–177.
21. Losick, R. & Pero, J. (1981) *Cell* **25**, 582–584.