MOLECULAR AND CELLULAR BIOLOGY, Sept. 1990, p. 4447–4455
0270-7306/90/094447-09$02.00/0

Vol. 10, No. 9

T_C, an Unusual Promoter Element Required for Constitutive Transcription of the Yeast *HIS3* Gene

SUBRAMONY MAHADEVAN† AND KEVIN STRUHL*

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

Received 12 April 1990/Accepted 30 May 1990

T_C is the proximal promoter element required for constitutive *his3* transcription that occurs in the absence of the canonical TATA element (T_R) and is initiated from the +1 site. The T_C element, unlike T_R, does not respond to transcriptional stimulation by the GCN4 or GAL4 activator protein. Analysis of deletion, substitution, and point mutations indicates that T_C mapped between nucleotides −54 and −83 and is a sequence-dependent element because it could not be functionally replaced by other DNA sequences. However, in contrast to the behavior of typical promoter elements, it was surprisingly difficult to eliminate T_C function by base pair substitutions. Of 15 derivatives averaging four substitutions in the T_C region and representing 40% of all possible single changes, only 1 inactivated the T_C element. Moreover, the phenotypes of mutant and hybrid elements indicated that inactivation of T_C required multiple changes. The spacing between T_C and the initiation region could be varied over a 30-base-pair range without significantly affecting the level of transcription from the +1 site. From these results, we consider it possible that T_C may not interact with TFIID or some other typical sequence-specific transcription factor, but instead might influence transcription, either directly or indirectly, by its DNA structure.

The promoter region of the *Saccharomyces cerevisiae his3* gene has been studied extensively in an attempt to elucidate the molecular mechanisms of eucaryotic transcriptional regulation (Fig. 1). During normal growth conditions, *his3* transcription occurs at a basal level (1 to 2 RNA molecules per cell at the steady state) and is initiated with equal efficiency from two sites, +1 and +13 (21). This constitutive mode of *his3* transcription requires a poly(dA) · poly(dT) sequence located about 120 base pairs (bp) upstream of the initiation region (20). However, when cells are subjected to conditions of amino acid starvation, transcription of *his3* is induced three- to fivefold over the basal level (26) along with many other amino acid-biosynthetic genes (reviewed in reference 9). This induction is mediated by the global regulator GCN4, which binds to a specific sequence located about 95 bp upstream of the *his3* initiation region (1, 10). Interestingly, GCN4-induced transcription is initiated with a strong preference for the +13 site; the level of +1 transcription is essentially unaffected (22). This pattern of inducible *his3* transcription is also observed in *gal-his3* hybrid promoters whose activity depends on GAL4 protein (19) and in promoters whose activity depends on *ope* suppressor mutations (12).

Mutational analysis has also identified two downstream promoter elements that are necessary for *his3* transcription (22). The T_R element is required for transcriptional activation by GCN4 or GAL4, and it can also support a low level of constitutive transcription dependent on the poly(dA) · poly(dT) element (7). T_R contains the sequence TATAAA and hence corresponds to a typical TATA element recognized by the TATA-binding protein TFIID. Almost all single-base-pair substitutions of TATAAA reduce T_R function, suggesting that the DNA sequence requirements for this element are fairly stringent (4, 7). Moreover, the activities conferred by these mutant TATA elements in vivo correlate very strongly with their levels of TFIID-dependent transcription in vitro, indicating that T_R functions by interacting with TFIID (28). *his3* transcription mediated by T_R is initiated almost exclusively from the +13 site (4, 7, 14, 22) because the element is too close to the +1 site for efficient initiation. Transcription from the +1 site is observed if the distance between T_R and the initiation region is increased even by as little as 8 bp (3, 14, 22).

The other downstream promoter element, T_C, has been inferred from the following observations (17, 22). First, small deletions (e.g., *his3*-Δ22; Fig. 1) that remove the T_R element do not significantly affect the basal level of *his3* expression even though they prevent GCN4 induction. Second, constitutive *his3* transcription occurs equally from the +1 and +13 sites even though T_R does not promote +1 initiation. Third, *his3*-Δ38, a deletion that removes sequences from −83 to −35, eliminates both constitutive and inducible *his3* expression. Thus, T_C is defined as the promoter-proximal element required for the basal level of transcription that occurs in the absence of T_R and is initiated from the +1 site. Taken together, the deletion analyses suggest that T_C includes sequences between −53 and −83. However, this localization must be regarded as preliminary as it relies on relatively large deletions, ignores the possibility of spacing effects between promoter elements, and is complicated by potential interactions (either positive or negative) between T_R and T_C.

Several lines of evidence suggest that T_C and T_R are functionally distinct elements and that T_C might not interact with TFIID (4, 7, 14, 22, 24). The region between −83 and −50 does not contain a sequence that resembles a typical TATA element; the CATAAT and AATGAA sequences most closely resembling the consensus contain two mutations (underlined), each of which significantly reduces the activity of the element. More importantly, the uninducibility of the +1 transcript strongly suggests that GCN4 and GAL4 are unable to activate transcription in combination with the

* Corresponding author.
† Present address: Molecular Biology Unit, Tata Institute of Fundamental Research, Homi Bhabha Road, Bombay 4OU 005, India.

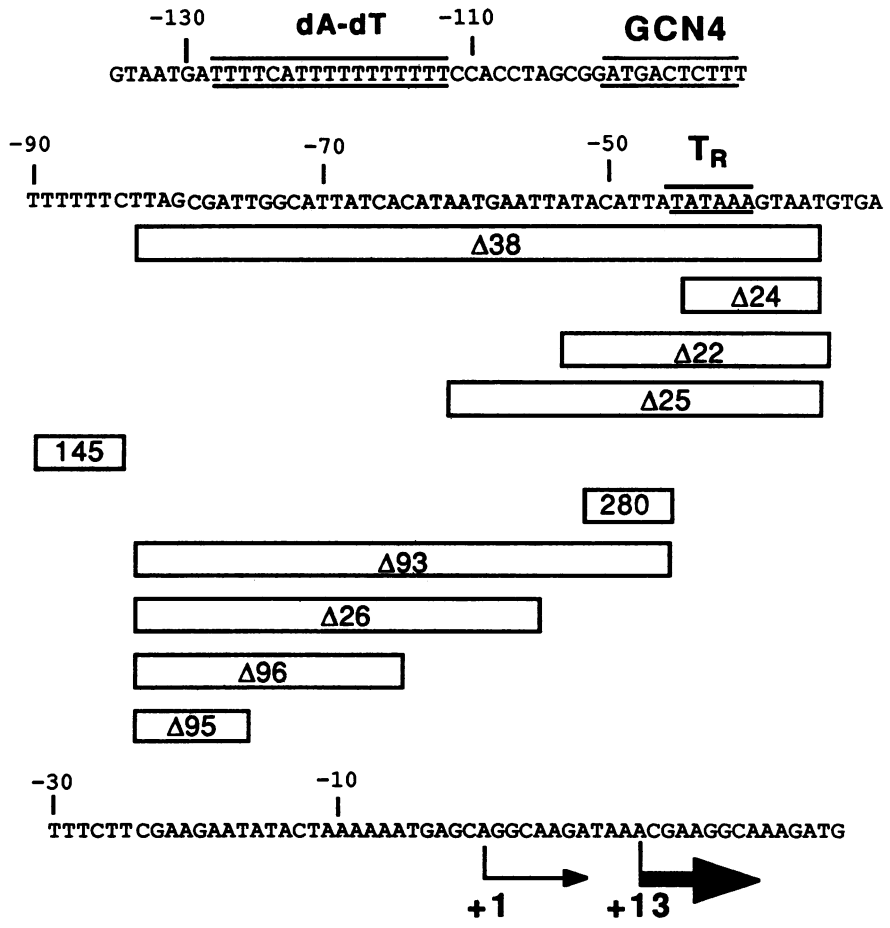

FIG. 1. Structures of the *his3* promoter region and selected deletion mutants. The DNA sequence from positions −135 to +26 (the AUG initiation codon for the *his3* gene product) is shown with the poly(dA) · poly(dT) element (20), GCN4-binding site (8, 10), and T_R element (4, 22, 28) over- and underlined and the +1 and +13 initiation sites indicated by arrows. Constitutive *his3* transcription initiates with equal efficiency from both sites, but GCN4-induced transcription occurs with a very strong preference for the +13 transcript (thick arrow) (14, 22). The extents of several deletion mutants are shown below the DNA sequence; *his3*-Δ22 (17), *his3-145* (8), *his3*-Δ93 (7), and *his3*-Δ38 and *his3*-Δ26 (18) have been described previously. The Δ38, Δ93, Δ26, Δ96, and Δ95 alleles all contain an *Eco*RI linker at the site of deletion, while *his3-145* and *his3-280* are linker-scanning mutations that respectively contain a *Sac*I or an *Eco*RI site.

T_C element. In this regard, GAL4 and GCN4 are also unable to induce transcription of the divergently transcribed *pet56* gene, presumably because its promoter also contains an "uninducible TATA element" (22). Finally, T_C-mediated transcription has not been observed in vitro with yeast nuclear extracts under experimental conditions that reproduce all known T_R-dependent phenomena (14). For these reasons, we have proposed that transcription mediated by the T_C and T_R elements occurs by different molecular mechanisms.

In this communication, we present a detailed mutational analysis of the *his3* T_C region. The results indicate that T_C is a sequence-dependent element that is surprisingly difficult to inactivate by single-base-pair substitutions. We consider it possible that T_C might not interact with a specific DNA-binding protein but instead might influence transcription by virtue of its DNA structure.

MATERIALS AND METHODS

DNA molecules. The starting molecules for most of the constructions were YIp55-Sc3740 and YIp55-Sc3745, both of which contain *his3*-Δ93, which removes sequences between −83 and −46 (7). Sc3740 carries the sequence TATAAA at the site of T_R flanked by *Eco*RI and *Sac*I linkers; Sc3745 is a similar molecule that carries the *his3-205* allele (TGTA AA). To facilitate DNA sequencing of the derivatives described below, we generated pTZ18-Sc4600, which contains the 1.5-kilobase (kb) *Sph*I-*Eco*RI fragment of YIp55-Sc3740 (ca. −1500 to −84) cloned into the pTZ18 vector (Pharmacia).

Derivatives containing point mutations in the T_C region were made by cloning a degenerate oligonucleotide averaging four substitutions per molecule. The single-stranded oligonucleotides were converted to double-stranded DNA by mutually primed synthesis (13), cleaved with *Eco*RI and *Dde*I, and ligated to *Bam*HI-*Eco*RI-cleaved pTZ18-Sc4600 and the 363-bp *Bam*HI-*Dde*I fragment (−447 to −80) of the wild-type *his3* promoter. Following DNA sequence analysis, the *Sph*I-*Eco*RI fragment of YIp55-Sc3745 was replaced by *Sph*I-*Eco*RI fragments containing the mutated T_C derivatives; thus, the resulting molecules contain the T_C derivatives in the context of the *his3-205* mutation that inactivates T_R. T_C derivatives containing the functional T_R element

were constructed analogously by replacing the *Sph*I-*Eco*RI fragment of YIp55-Sc3740.

Hybrid T_C elements and "filler" mutations were constructed by using fortuitous restriction sites created in the mutated derivatives described above. The *his3-275* allele was made by blunting the *Mlu*I site in *his3-255* and sequentially inserting *Eco*RI and *Kpn*I linkers; the resulting molecule has a substitution of 11 bp at the 3′ end of T_C. DNA containing *his3-276* was made by replacing the 12 bp from the center of the *Bst*UI in *his3-265* up to the *Eco*RI site with the 13-bp *Sma*I-*Eco*RI fragment from the pUC18 polylinker.

The internal deletions *his3*-Δ26, *his3*-Δ95, and *his3*-Δ96 were constructed by fusing *his3* sequences from Sc3284, Sc3285, and Sc3286 (endpoints −55, −65, and −75, respectively; Struhl, unpublished) downstream of the *Eco*RI site at −84 in YIp55-Sc3745. Deletions that reduce the spacing between T_C and the *his3* initiation region (*his3*-Δ97, *his3*-Δ98, and *his3*-Δ99) were constructed by fusing *his3* sequences from Sc3704, Sc3705, and Sc3706 (endpoints −24, −35, and −31, respectively) (3) downstream of the *Eco*RI site at −52 present in YIp55-Sc4640 (contains the wild-type T_C element) or YIp55-Sc4645 (contains the *his3-250* allele). The *his3*-Δ100 insertion mutation that alters the spacing between T_C and the initiation region was obtained by introducing a *Sal*I linker at the *Eco*RI site in YIp55-Sc4640 or YIp55-Sc4645.

Phenotypic analysis. YIp55 DNA molecules containing the various T_C and T_R alleles were introduced into strain KY320 by precise replacement of the *his3* chromosomal location as described previously (21). The resulting derivatives were examined for their 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 (4, 7, 8, 23). Strains carrying the wild-type T_C element but deleted or mutated for T_R can grow well in medium containing 20 mM AT. Phenotypes of T_C derivatives were scored as follows: +++, normal growth on 20 mM AT; ++, normal growth on 10 mM AT, weak growth on 20 mM AT; +, normal growth on 5 mM AT, weak growth on 10 mM AT; +/−, weak growth on 5 mM AT; −, no growth on 5 mM AT. For some experiments, *his3* deletion alleles were introduced by gene replacement into AY824 (*gcn4-Δ1 bas2-2 ura3-52*; kindly obtained from Kim Arndt) or KY329 (relevant genotype: *ura3-52 his3-TRP1 gcn4-Δ1*) (23).

To analyze *his3* transcription directly, total RNAs from cells grown under normal conditions (30°C in YPD broth) were hybridized to completion with an excess of ^{32}P-end-labeled *his3* and *ded1* oligonucleotide probes and treated with S1 nuclease. The products were separated by electrophoresis in 10% denaturing polyacrylamide gels and visualized by autoradiography (4).

RESULTS

Transcriptional initiation pattern mediated by T_C. As mentioned in the Introduction, T_R-mediated transcription is initiated almost exclusively from the +13 site (4, 7, 14, 25). However, the transcriptional contribution of T_C was difficult to determine because the derivatives examined either contained a functional T_R element or had an altered spacing between T_C and the initiation region. To circumvent these problems, we examined the transcription pattern from *his3-205*, a derivative that contains a point mutation of the T_R element (TGTAAA) but retains the T_C element at its normal position in the *his3* promoter. As shown in Fig. 2, the

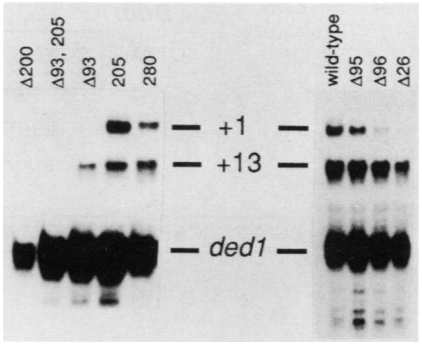

FIG. 2. Initiation pattern mediated by the T_C and T_R elements. S1 analysis of RNA from the indicated derivatives (see Fig. 1), including Δ200, which completely deletes the *his3* promoter region and structural gene. The positions of the +1 and +13 *his3* RNAs and the *ded1* RNA are indicated.

his3-205 strain expressed wild-type levels of the +1 transcript but reduced levels of the +13 transcript. This pattern depended on the T_C element, because a strain which carried a deletion of sequences from −83 to −55 as well as the T_R point mutation (*his3*-Δ93 *his3-205*) showed essentially no transcription from either the +1 or the +13 site. As expected (7), a *his3*-Δ93 strain that carried the same T_C deletion but a normal T_R sequence (TATAAA) expressed the +13 transcript almost exclusively. Finally, equal levels of the +1 and +13 transcripts were observed in the *his3-280* strain, which contains an *Eco*RI linker substitution between normal T_C and T_R sequences. Therefore, the T_C element is essential for initiation from the +1 site, and it, along with the T_R element, contributes to constitutive transcription from the +13 site.

Deletions of the T_C region. The fact that the normal level of +1 transcription was conferred by *his3-280*, in which the region between −46 and −53 was replaced by an *Eco*RI linker, defined the 3′ boundary of T_C at −54. This boundary is consistent with the phenotype of *his3*-Δ22, which removed sequences between −53 and −35 yet retained basal level expression (17) (see Fig. 1 and later). The 5′ boundary can be placed at position −83 by the previous observation that a *Sac*I linker between −84 and −90 (*his3-145*) did not affect the basal level of transcription (8) (Fig. 1). To define the 5′ boundary of T_C more precisely, three internal deletions with a common upstream point at −83 and extending downstream to −75, −65, and −55, with *Eco*RI linkers at the junctions, were analyzed (Fig. 1 and 2). In *his3*-Δ95, in which 8 bp at the distal end of T_C were replaced precisely by the *Eco*RI linker, transcription from the +1 site was slightly reduced in comparison to transcription from the +13 site. In the other two cases (*his3*-Δ96 and *his3*-Δ26), the +1 transcript was initiated inefficiently, with the least amount being detected in the case of the largest deletion. Thus, the T_C element requires sequences between −65 and −75, and sequences between −75 and −83 can affect T_C function.

T_C point mutants. By analogy with numerous other promoter elements, a likely role for the T_C element in mediating +1 transcription is by acting as a protein-binding site. From the properties of a wide variety of specific DNA-binding transcription factors, one would expect T_C to be a highly specific recognition sequence whose transcriptional activity is very sensitive to point mutations. Indeed, detailed mutagenesis of two *his3* promoter sequences, the GCN4 binding site (8) and the T_R element (4, 7, 28), indicates that most single-base-pair substitutions significantly reduce DNA binding of the transcription factor and promoter activity in

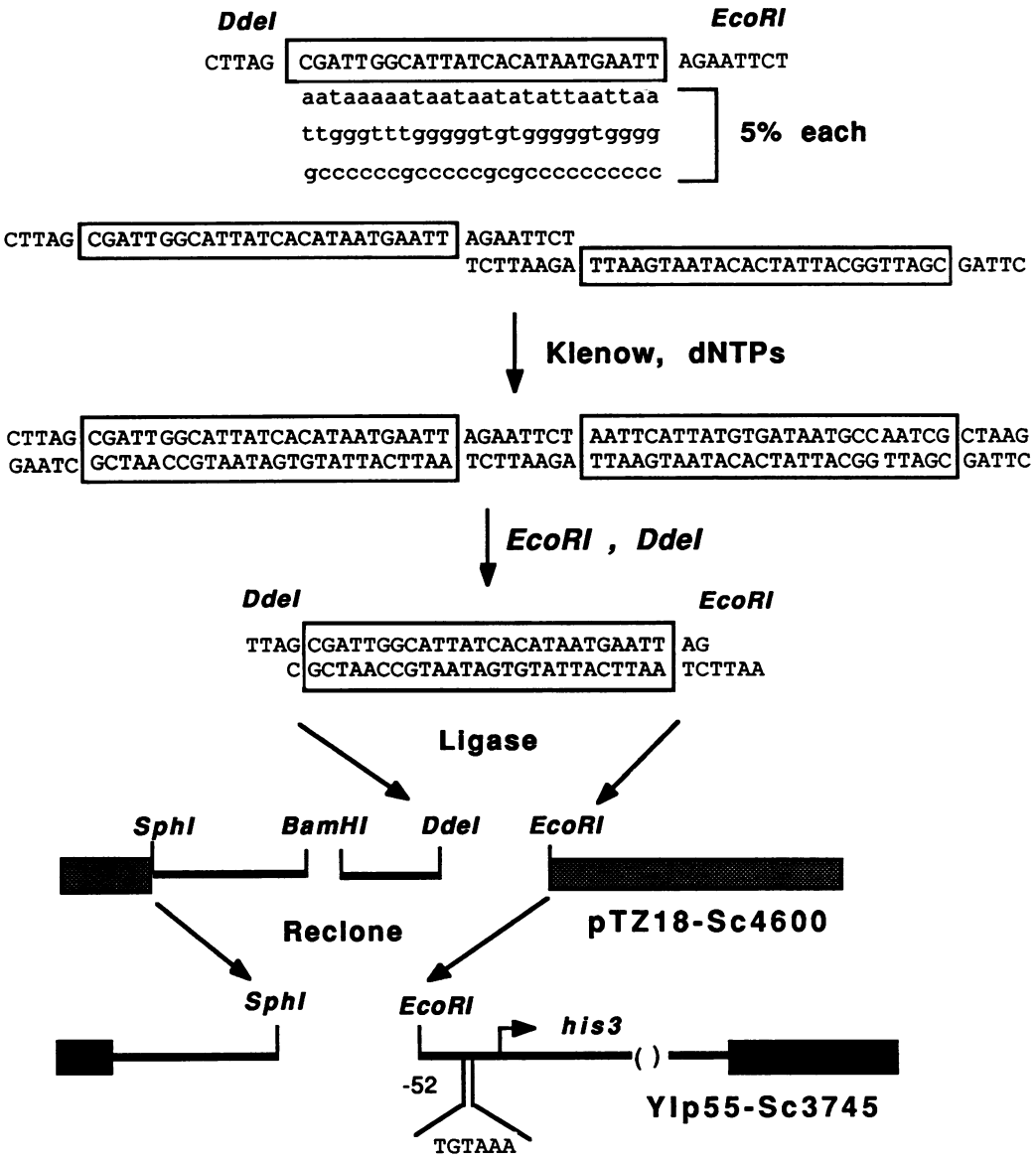

FIG. 3. Construction of mutations in the *his3* T_C element. A degenerate oligonucleotide containing the T_C region (boxed) flanked by *Dde*I and *Eco*RI sites at the 5′ and 3′ ends, respectively, was synthesized with 5% of each non-wild-type nucleotide (lowercase letters) at the relevant positions. This mixture of single-stranded oligonucleotides was converted to double-stranded *Dde*I-*Eco*RI fragments by the mutually primed synthesis procedure (13) and cloned into pTZ18-Sc4600, which contains the 1.5-kb *Sph*I-*Eco*RI fragment corresponding to positions ca. −1500 to −84. The resulting *Sph*I-*Eco*RI fragments were then cloned upstream of the mutated T_R element in YIp55-Sc3745.

vivo. Although the T_C region does not contain a sequence that resembles a consensus TATA element, it does contain a dyad-symmetric sequence between −72 and −58 in which inverted copies of CATTAT are separated by 3 bp. In addition, a related sequence, CATΔAT, is found between −64 and −59.

To examine the DNA sequence requirements for T_C function, we analyzed the phenotypes of mutated T_C derivatives in the context of the *his3-205* mutation, which inactivates the T_R element; as shown in Fig. 2, *his3* transcription in this situation depended on T_C. Specifically, a library of T_C mutants was created by cloning a degenerate oligonucleotide averaging four mutations per molecule corresponding to positions −79 to −54 between the relevant *Dde*I and *Eco*RI sites of *his3-205* DNA (Fig. 3). The resulting molecules were introduced into the yeast genome by gene replacement, and the effect of the mutations on T_C function was determined by monitoring growth in the presence of AT and by measuring *his3* RNA levels (Fig. 4).

Surprisingly, 14 out of the 16 derivatives tested, including those with as many as five or six point mutations, retained T_C function. In many cases, the growth and transcriptional properties of these derivatives were indistinguishable from those conferred by the wild-type T_C element; in some cases, the derivatives were somewhat less active. Several of the fully functional T_C derivatives significantly disrupted the dyad-symmetric sequence between −72 and −58. These included *his3-260*, with five mutations spread over both sides of the dyad; *his3-268*, with two mutations in the left half of the dyad and three mutations between the halves; and

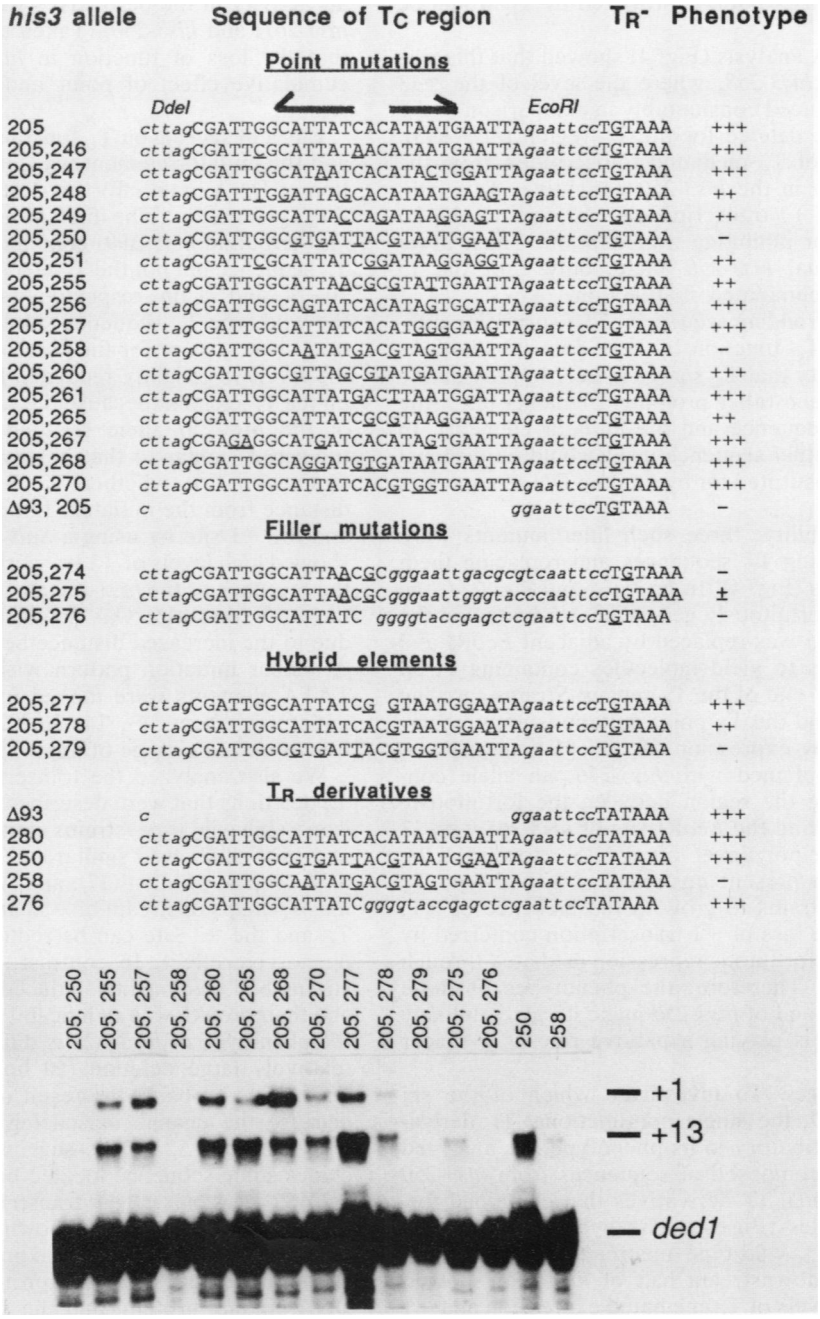

FIG. 4. Sequences and phenotypes of the T_C mutations. The sequences of the various derivatives are shown, with base pair substitutions underlined and linker substitutions indicated by lowercase italic letters (the *Dde*I site is naturally occurring). The dyad-symmetric sequence in the T_C region is indicated by a two-headed arrow. The phenotypes were defined by growth in medium containing AT as described in Materials and Methods. All of the alleles have been examined in combination with a mutated T_R element (*his3-205*), and some of them have been examined in the context of a wild-type T_R element. Shown below are the results of S1 analysis of RNA from the indicated derivatives, with the positions of the +1 and +13 *his3* RNAs and the *ded1* RNA indicated. We do not understand the basis for the apparent variability in the relative amounts of the +1 and +13 RNAs that was observed in certain alleles.

his3-257 and *his3-270*, which both have three mutations in the downstream half of the dyad. Taken together, these 14 functional T_C derivatives contain 31 out of the 78 possible single-base-pair changes.

Only two mutations, *his3-250* and *his3-258*, appeared to inactivate the T_C element; strains containing these derivatives failed to grow in AT and did not express the +1 transcript. However, since most of the derivatives with multiple mutations did not appreciably affect T_C function, it is conceivable that the phenotypes of these two exceptional derivatives might be due to the creation of a fortuitous repressor binding site, resulting in inhibition of all *his3* transcription. To test this possibility, we replaced the defective T_R allele (*his3-205*) with the wild-type T_R sequence at its normal location downstream of the *his3-250* and *his3-258* alleles. In the event of a repressor occupying the T_C region,

reduced levels of transcription contributed by T_R would be expected.

Interestingly, RNA analysis (Fig. 4) showed that this was indeed the case for *his3-258*, where the level of the +13 transcription was reduced considerably in comparison to the equivalent derivative deleted for the T_C region (*his3*-Δ93). However, the level of T_R-mediated transcription from the +13 site was normal in the *his3-250* derivative. This indicates that the loss of +1 transcript in *his3-250* is not due to a fortuitous repressor inhibiting transcription. Thus, if we exclude the artifactual *his3-258* allele, only 1 of the 15 mutated derivatives eliminated T_C function.

Replacing T_C with random sequences. The surprising difficulty in eliminating T_C function by base pair substitutions suggests the possibility that T_C serves merely to provide the proper spacing between other promoter elements [e.g., the poly(dA) ·poly(dT) sequence and the initiator element]. In such a model, the actual sequence itself would be dispensable and could be substituted for by random DNA sequences of the same length.

To test this possibility, three such filler mutants were constructed by deleting T_C sequences and replacing them with linker sequences (Fig. 4). In *his3-275* and *his3-274*, the region between the fortuitously generated *Mlu*I site and the *Eco*RI site of *his3-255* was replaced by adjacent *Eco*RI and *Kpn*I or *Mlu*I linkers to yield molecules containing 11-bp substitutions at the 3′ end of the T_C region. Strains carrying these substitutions and the T_R point mutations grew poorly in AT and had a low expression of the +1 transcript. A similar result was obtained with *his3-276*, an allele constructed by replacing the region between the fortuitously generated *Bst*UI site and the *Eco*RI site of *his3-265* with 13 nucleotides from the polylinker of pUC18. However, the *his3-276* allele, when present upstream of a wild-type T_R element, permitted strains to grow in the presence of AT; this indicates that the loss of +1 transcription conferred by *his3-276* is not due to fortuitous repression mediated through the linker sequence. Therefore, the phenotypes of these three filler mutations and of *his3-250* make it highly unlikely that the T_C element is playing a passive role as a spacer sequence.

Hybrid T_C sequences. To investigate which of the six mutations in *his3-250*, the single nonfunctional T_C derivative, were most contributory to its phenotype, we analyzed hybrid T_C elements composed of sequences from *his3-250* and from two functional T_C derivatives that contained fortuitous restriction sites (Fig. 4). The hybrid sequence in *his3-278*, which has a wild-type upstream half and point mutations from the downstream half of *his3-250*, showed somewhat reduced levels of T_C-mediated expression and +1 transcription. A similar derivative, *his3-277*, which has the nearly wild-type upstream half from *his3-265*, showed full T_C function. The wild-type phenotype of *his3-277* is also noteworthy because this derivative contains a single-base-pair deletion in the center of the T_C region. Therefore, the point mutations present within the downstream half of *his3-250* are not sufficient to inactivate the element.

Conversely, strains carrying *his3-279*, which has the mutations from the upstream half of *his3-250* and the downstream half of *his3-270*, showed a loss of T_C-dependent transcription from the +1 site. This suggests that the three mutations at positions −71, −69, and −66 in *his3-250* are primarily responsible for its phenotype. However, the A to G change at −71 and the T to G change at −69 were present in fully functional T_C derivatives (*his3-260*, *his3-267*, and *his3-268*), and a C to G transversion at −66 did not influence T_C activity in the derivatives where it was found (*his3-258*, *his3-261*, and *his3-268*). Taken together, the results suggest that the loss of function in *his3-250* is the result of the cumulative effect of point mutations spread over the T_C region.

Effect of spacing on T_C function. The distance between T_C and the initiator element was altered by deletion and insertion of DNA in an effort to detect any spacing requirement for T_C function. The first group of derivatives *his3*-Δ97, *his3*-Δ98, and *his3*-Δ99) contained the wild-type or *his3-250* T_C sequence but not the T_R element, resulting in deletions of 10, 14, and 21 bp, respectively (Fig. 5). Deletions carrying the wild-type T_C sequence showed normal levels of the +1 transcript, except for the case of the 21-bp deletion, where the level was slightly reduced. This transcription depended on the T_C element because the same deletions in the context of the *his3-250* allele resulted in low levels of the +1 transcript similar to that observed in the parental *his3-250* strain. To test whether T_C could function at increased distance from the initiator, 12 bp were inserted between T_C and the +1 site by using a *Sal*I linker. This derivative also showed high levels of +1 transcription that was eliminated in the presence of the *his3-250* allele. Interestingly, the level of the +13 transcript was significantly reduced, presumably due to the increased distance between T_C and the +13 site. A similar initiation pattern was observed when canonical TATA elements were moved further upstream of the *his3* initiator elements (3, 14, 16, 22). Thus, T_C is capable of function over a range of approximately 30 nucleotides.

We also analyzed the transcriptional properties of small T_R deletions that were described previously (17). In comparison to the wild type, strains containing *his3*-Δ20, -Δ22, -Δ23, and -Δ24 all showed similar levels of the +1 transcript (Fig. 6), *his3* gene product (17), and AT resistance. This suggests that T_C maps upstream of −53 and that the spacing between T_C and the +1 site can be reduced by 23 bp with minimal phenotypic effect. In contrast, the *his3*-Δ25 and *his3*-Δ21 strains had significantly reduced levels of +1 transcription and their growth was significantly inhibited by AT. Although the phenotype of *his3*-Δ21 is difficult to interpret due to the relatively large deletion (50 bp), the initiation pattern of *his3*-Δ25 is probably the result of removal of T_C sequences because the change in spacing is very similar to that of *his3*-Δ22 and -Δ23. This suggests that the T_C element includes some sequence located between −54 and −62.

***BAS2* does not affect transcription mediated by the T_C element.** The *BAS2* (also known as *PHO2*) gene product is essential for the basal-level expression of *HIS4*, a gene that is also activated by GCN4 protein (2). BAS2 is a specific DNA-binding protein, and the sequences of several target sites are known (2). The dyad-symmetric sequence in the T_C region resembles the BAS2-binding sites, and BAS2 activation of *his4* transcription does not require the TATA element necessary for GCN4 activation (2). To investigate the possibility that BAS2 might be the T_C-binding protein responsible for constitutive *his3* expression, the phenotypes of the T_R deletion mutants were analyzed in a *bas2* mutant strain. The level and initiation pattern of *his3* transcription conferred by these T_R deletion alleles were unaffected in the *bas2* strain (Fig. 6), indicating that BAS2 does not affect T_C-mediated transcription.

DISCUSSION

Properties of the T_C element. The above experiments demonstrate that T_C is necessary for constitutive *his3* tran-

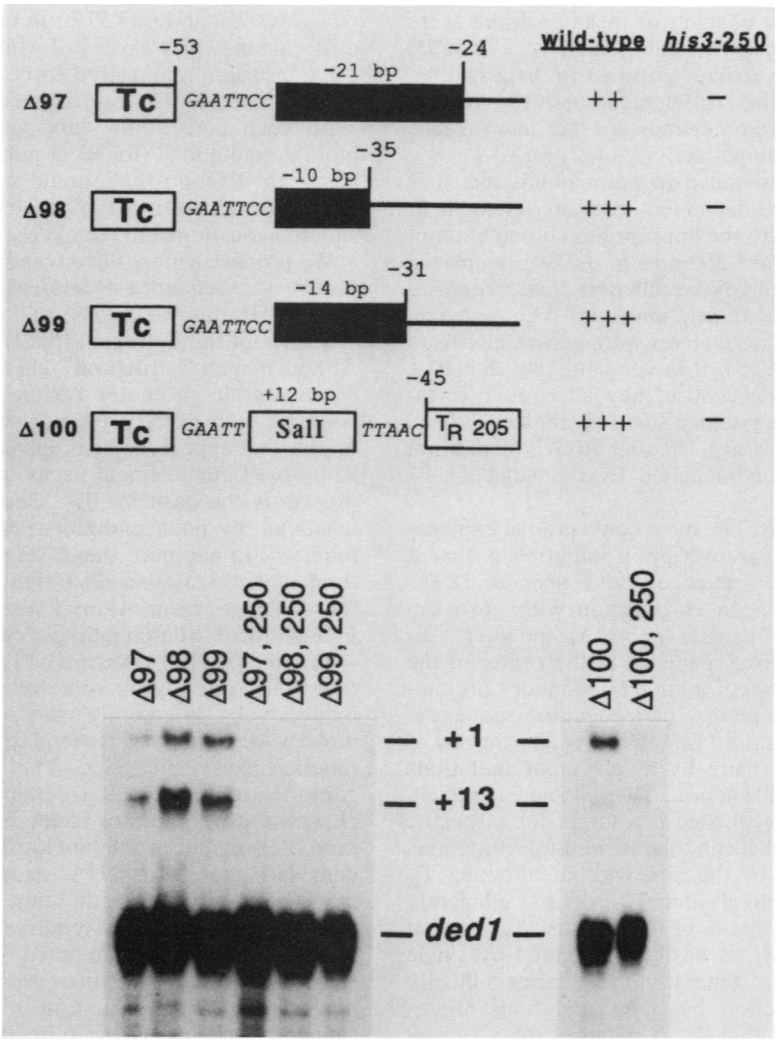

FIG. 5. Analysis of spacing mutations. The structures of the various derivatives are shown, with deletions indicated by solid boxes, the *SalI* linker insertion indicated by an open box, linker substitutions indicated by italic letters, and the total spacing change shown above the insertion or deletion. The phenotypes, in combination with the wild-type and mutated (*his3-250*) T_C element, were defined by growth in medium containing AT as described in Materials and Methods. Shown below are the results of S1 analysis of RNA from the indicated derivatives, with the position of the +1 and +13 *his3* RNAs and the *ded1* RNA indicated.

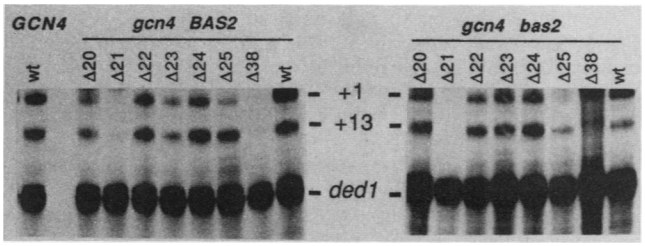

FIG. 6. Effect of BAS2 on T_R deletion mutations. S1 analysis of RNA from *BAS2* or *bas2* strains containing the indicated *his3* alleles; in all cases except for the lane at the extreme left, the strains also contain the *gcn4*-Δ1 deletion. The endpoints of the deletion mutations have been determined previously (17) and are as follows: Δ20, −46 to −34; Δ21, −59 to −9; Δ22, −53 to −34; Δ23, −50 to −27; Δ24, −44 to −35; Δ25, −35 to −61; Δ38, −83 to −35. The positions of the +1 and +13 *his3* RNAs and the *ded1* RNA are indicated.

scription initiated from the +1 site and is contributory to basal-level expression from the +13 site. Transcription mediated by T_C does not require the consensus TATA element, T_R, that interacts with the TATA-binding factor TFIID (28). The region between −83 and −53 is necessary and sufficient to encode the T_C element. The distance between T_C and the *his3* initiation region can be varied over a 30-bp range with only minimal effects on the level of transcription from the +1 initiation site.

The most surprising observation about T_C is that its activity is very refractory to single-base-pair substitutions. Excluding *his3-258*, which fortuitously created an inhibitory element, the 15 derivatives in Fig. 4 contain an average of four mutations in a 26-bp region and represent 40% of all possible substitutions. Nevertheless, only *his3-250*, which contained six substitutions spread over the region, eliminated T_C function. Even in this exceptional case, multiple mutations were necessary for the phenotype. The phenotypes conferred by the hybrid T_C elements suggest that the mutations at positions −71, −69, and −66 are most impor-

tant, but identical or other changes at these positions were found in several derivatives with full T_C activity (*his3-258*, *his3-261*, and *his3-268*). In striking contrast to these results, similar mutageneses of the T_R element and the GCN4-binding site in the *his3* promoter indicate that many single substitutions abolish functional activity (4, 7, 8).

Despite its apparent resistance to point mutations, it is clear that T_C is a sequence-dependent promoter element. It can be inactivated either by the appropriate combination of base pair substitutions (*his3-250* and *his3-279*), by partial replacement with unrelated DNA sequences (*his3-274*, *his3-275*, and *his3-276*), or by small deletions (*his3*-Δ25, *his3*-Δ26, and *his3*-Δ96). Since these derivatives were generated before their phenotype was assessed, it is very unlikely that they fail to confer T_C function because they all contain fortuitously created negative regulatory sites. Moreover, at least two defective T_C alleles (*his3-250* and *his3-276*) did not interfere with transcription mediated by the adjacent T_R element.

Molecular interpretations. The most conventional explanation for the role of T_C in transcription initiation is that it provides the recognition sequence for a specific DNA-binding protein that works in conjunction with upstream promoter elements such as the poly(dA) · poly(dT) sequence. The dyad-symmetric sequence in the center of the T_C region is a tempting target for such a regulatory protein. However, a characteristic feature of recognition sequences for specific transcription factors is mutability; altering any of the 5 to 10 crucial base pairs by single point mutations generally results in loss of function. Thus, it seems unlikely that the dyad-symmetric sequence is a target for a specific T_C-binding protein because it can tolerate multiple mutations in either or both halves of the site without affecting T_C function. It also seems unlikely that T_C interacts efficiently with the TATA-binding protein TFIID because T_R, a good TFIID interaction site (28), is easily inactivated by single point mutations (4, 7). More generally, the extreme difficulty in knocking out T_C function by point mutations argues against the simple model that T_C is a target for a typical transcription factor.

The surprising mutagenesis results could be explained by proposing multiple elements within the T_C region, any one of which would be sufficient to support transcription. By such a model, T_C function could be eliminated only by inactivating all of these putative elements. Such redundancy is often found in eucaryotic promoters and could reflect contiguous or overlapping sites within the T_C region that recognize the same regulatory protein or different proteins. It is also possible that redundancy could reflect multiple weak TFIID interaction sites, although it would be difficult to explain why T_C is transcriptionally inactive in vitro (14), unable to respond to upstream activator proteins (22), and does not show detectable DNase I protection by purified TFIID (Ponticelli and Struhl, unpublished data).

Although the possibility of redundant elements cannot be excluded, the mutational analysis puts some constraints on such a hypothesis. The fact that *his3*-Δ25 and *his3*-Δ96 inactivated T_C indicates that (i) a functional element cannot exist between −83 and −62 or between −53 and −65 and (ii) neither half-dyad is sufficient for T_C function. Thus, redundant elements would have to overlap in the region of the dyad-symmetric sequence. Moreover, the region of redundancy would have to be fairly large to account for why T_C function requires a sequence between −54 and −62 (compare the phenotypes of *his3*-Δ25 and *his3*-Δ22) yet can be inactivated by a combination of mutations at −71, −69, and −66 (*his3-250* and *his3-279*). In fact, sequences even further upstream may play a role in T_C function, because the level of +1 transcription appeared to be somewhat reduced by the *Eco*RI linker substitution between −75 and −83 (*his3*-Δc). With such constraints, however, one might expect that multiple mutations (found in many derivatives) or a single-base-pair deletion that should result in a major structural distortion of the dyad-symmetry region (*his3-277*) would simultaneously inactivate overlapping elements.

We propose a more unconventional model for T_C, namely, that it is a sequence-dependent but not sequence-specific element. In this view, it is the structure but not the precise sequence of the T_C region that is important for its function. Although such "structural" elements are not characteristic of eucaryotic promoter regions, they do occur in other contexts, such as the A+T-rich sequence in DNA replication origins that appears to be involved in unwinding (27) and the 80-bp A+T-rich element necessary for centromere function (6). As is the case for T_C, these elements are difficult to inactivate by point mutations; more extensive changes are required. In addition, the T_C region, like these other structural elements, is also A+T rich; however, derivatives with four changes from A or T residues to G (*his3-257* and *his3-260*) had little phenotypic consequence.

In principle, the structure of T_C could influence transcription either directly by some inherent physical property or indirectly by interacting with a protein. Whether it acts directly or indirectly, several specific mechanisms for T_C function can be imagined. The T_C region could assume a conformation that facilitates interaction of RNA polymerase II with the +1 initiator element, facilitates melting of the promoter region, or inhibits local nucleosome formation. In considering such a DNA structure-based transcriptional mechanism, it is interesting that efficient T_C-mediated initiation from the +1 site requires a poly(dA) · poly(dT) sequence (20) and is associated with an altered chromatin structure (evidenced by increased sensitivity to micrococcal nuclease) in the region at or near T_C (22). Poly(dA) ·poly(dT) sequences are inhibitory to nucleosome formation (11, 15), and they can stimulate transcription by T7 RNA polymerase and RNA polymerase II to comparable extents in yeast cells (5). Thus, the unusual nature of the T_C element is consistent with previous suggestions that constitutive *his3* transcription from the +1 site may involve effects on the chromatin template (5, 20, 22, 24) and may not require binding of TFIID to the promoter (4, 7, 14, 22, 28).

ACKNOWLEDGMENTS

We thank Wei Chen for performing the RNA analysis shown in Fig. 6, Kim Arndt for strain AY824, and Fred Ponticelli for useful comments during the course of the work and on the manuscript.

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

LITERATURE CITED

1. **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.
2. **Arndt, K. T., C. Styles, and G. R. Fink.** 1987. Multiple global regulators control *HIS4* transcription in yeast. Science **237:** 874–880.
3. **Chen, W., and K. Struhl.** 1985. Yeast mRNA initiation sites are determined primarily by specific sequences, not by the distance from the TATA element. EMBO J. **4:**3273–3280.
4. **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.
5. **Chen, W., S. Tabor, and K. Struhl.** 1987. Distinguishing between mechanisms of eukaryotic transcriptional activation with bacteriophage T7 RNA polymerase. Cell **50**:1047–1055.
6. **Gaudet, A., and M. Fitzgerald-Hayes.** 1987. Alterations in the adenine-plus-thymine-rich region of *CEN3* affect centromere function in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **7**:68–75.
7. **Harbury, P. A. B., and K. Struhl.** 1989. Functional distinctions between yeast TATA elements. Mol. Cell. Biol. **9**:5298–5304.
8. **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.
9. **Hinnebusch, A. G.** 1986. The general control of amino acid biosynthetic genes in the yeast *Saccharomyces cerevisiae*. Crit. Rev. Biochem. **21**:277–317.
10. **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.
11. **Kunkel, G. R., and H. G. Martinson.** 1981. Nucleosomes will not form on double stranded RNA or over poly(dA):(dT) tracts in recombinant DNA. Nucleic Acids Res. **9**:6869–6888.
12. **Oettinger, M. A., and K. Struhl.** 1985. Suppressors of *Saccharomyces cerevisiae* promoter mutations lacking the upstream element. Mol. Cell. Biol. **5**:1901–1909.
13. **Oliphant, A. R., A. L. Nussbaum, and K. Struhl.** 1986. Cloning of random-sequence oligodeoxynucleotides. Gene **44**:177–183.
14. **Ponticelli, A. S., and K. Struhl.** 1990. Analysis of yeast *his3* transcription in vitro: biochemical support for multiple mechanisms of transcription. Mol. Cell. Biol. **10**:2832–2839.
15. **Prunell, A.** 1982. Nucleosome reconstitution on plasmid inserted poly(dA:dT). EMBO J. **1**:173–179.
16. **Singer, V. L., C. R. Wobbe, and K. Struhl.** 1990. A wide variety of DNA sequences can functionally replace a yeast TATA element for transcriptional activation. Genes Dev. **4**:636–645.
17. **Struhl, K.** 1982. Regulatory sites for *his3* expression in yeast. Nature (London) **300**:284–287.
18. **Struhl, K.** 1982. The yeast *his3* promoter contains at least two distinct elements. Proc. Natl. Acad. Sci. USA **79**:7385–7389.
19. **Struhl, K.** 1984. Genetic properties and chromatin structure of the yeast *gal* regulatory element; an enhancer-like sequence. Proc. Natl. Acad. Sci. USA **81**:7865–7869.
20. **Struhl, K.** 1985. Naturally occurring poly(dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. Proc. Natl. Acad. Sci. USA **82**:8419–8423.
21. **Struhl, K.** 1985. Nucleotide sequence and transcriptional mapping of the yeast *pet56-his3-ded1* gene region. Nucleic Acids Res. **13**:8587–8601.
22. **Struhl, K.** 1986. Constitutive and inducible *Saccharomyces cerevisiae* promoters: evidence for two distinct molecular mechanisms. Mol. Cell. Biol. **6**:3847–3853.
23. **Struhl, K.** 1987. The DNA-binding domains of the Jun oncoprotein and the yeast GCN4 transcriptional activator are functionally homologous. Cell **50**:841–846.
24. **Struhl, K.** 1987. Promoters, activator proteins, and the mechanism of transcriptional initiation in yeast. Cell **49**:295–297.
25. **Struhl, K., W. Chen, D. E. Hill, I. A. Hope, and M. A. Oettinger.** 1985. Constitutive and coordinately regulated transcription of yeast genes: promoter elements, positive and negative regulatory sites, and DNA-binding proteins. Cold Spring Harbor Symp. Quant. Biol. **50**:489–503.
26. **Struhl, K., and R. W. Davis.** 1981. Transcription of the *his3* gene region in *Saccharomyces cerevisiae*. J. Mol. Biol. **152**:535–552.
27. **Umek, R. M., and D. Kowalski.** 1988. The ease of unwinding as a determinant of initiation of yeast replication origins. Cell **52**:559–567.
28. **Wobbe, C. R., and K. Struhl.** 1990. Yeast and mammalian TATA-binding proteins have nearly identical DNA sequence requirements for transcription in vitro. Mol. Cell. Biol. **10**:3859–3867.