Mutational analysis of the transcription start site of the yeast tRNA$^{\text{Leu}}_3$ gene

Paolo Fruscoloni$^{1,2}$, Michela Zamboni$^3$, Gianna Panetta$^1$, and Augusto De Paolis and Glauco P. Tocchini-Valentin$^{1,2,}$

$^1$Institute of Cell Biology, CNR, Viale Marx 43, 00137 Rome, Italy, $^2$Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL 60637, USA and $^3$EniChem, Istituto Donegani, 00015 Monterotondo, Rome, Italy

Received April 20, 1995; Revised and Accepted June 20, 1995

ABSTRACT

In addition to the well-known internal promoter elements of tRNA genes, 5' flanking sequences can also influence the efficiency of transcription by Saccharomyces cerevisiae extracts in vitro. A consensus sequence of yeast tRNA genes in the vicinity of the transcriptional start site can be derived. To determine whether the activity of this region can be attributed to particular sequence features we studied in vitro mutants of the start site region. We found that the start site can be shifted, but only to a limited extent, by moving the conserved sequence element. We found that both a pyrimidine-purine motif (with transcription initiating at the purine) and a small T:A base pair block upstream are important for efficient transcription in vitro. Thus the sequence surrounding the start site of transcription of the yeast tRNA$^{\text{Leu}}_3$ gene does play a role in determining transcription efficiency and fixing the precise site of initiation by RNA polymerase III.

INTRODUCTION

The 5' flanking region is generally of fundamental importance in the transcription of genes in prokaryotes and eukaryotes. All three eukaryotic nuclear RNA polymerases require the TATA binding protein (TBP) for transcription (1-4). The mechanisms of action of TBP are best understood for the 5' flanking sequences of class II nuclear genes, the genes that code for proteins. The 8bp of the TATA box bind to the concave surface of TBP by bending towards the major groove; as a result of binding a wide, underwound, shallow minor groove is produced. However, the role of the 5’ flanking region of class III genes is much less understood, since most attention has so far been payed only to the major elements of the promoter, which are intragenic. Transcription of eukaryotic tRNA genes is generally dependent upon two highly conserved intragenic sequence elements, the A and B boxes, which function as binding sites for the RNA–Pol III-specific transcription factor TFIIIC.

The influence of 5’ flanking regions on the transcription of tRNA genes has been demonstrated in vitro (5-7, see 8 for a review). It has been observed that in most cases deletions or substitutions in the 5’ flanking regions diminish and, in some instances, almost preclude transcription, while in a few instances an enhancing effect has been found (9). Effects are strikingly position-dependent; changes in the position of portions of the region by only a few base pairs relative to the coding sequence can significantly alter transcriptional efficiency. The importance of 5’ flanking sequences on transcription manifests itself differently in cell-free systems derived from different organisms (8).

Yeast transcription factor TFIIIB, a complex of TBP with two other proteins, is the central transcription factor of RNA polymerase III (Pol III). It is both necessary and sufficient to direct Pol III to the transcriptional start site of tRNA genes (10). TFIIIB appears not to recognize specific sequences, but to be directed to its 40 bp long DNA binding site, upstream of the transcriptional start site, by intragenically bound TFIIIC (11). Pol III is recruited to the promoter by TFIIIB. Once there, Pol III can form an open transcription complex (12). The selection of the start site and the efficiency of transcription of tRNA genes are directed by an initiation window established by TFIIIB, which directs the catalytic subunits of Pol III toward the window. The primary tRNA transcripts contain 5’ and 3’-ends that are extended and, occasionally, also a small intervening sequence. These tRNA precursors are processed to mature sized tRNA by nucleases such as RNase P and the splicing endonuclease.

The primary aim of the experiments described in this paper was a detailed investigation of the nature of the contribution of 5’ flanking sequences to in vitro transcription initiation of tRNA genes by Pol III. Particular conditions of in vitro transcription were used to produce unprocessed primary transcripts (see Materials and Methods). It was found that the sequences surrounding the start site of tRNA gene transcription play a specific role in both determination of transcriptional efficiency and in fixing the precise site of initiation.

MATERIALS AND METHODS

Plasmids and mutagenesis

Plasmid pTGE2M was constructed by subcloning the 165 bp S$\text{al}$ DNA fragment containing the tRNA$^{\text{Leu}}_3$ gene from plasmid pTO (13) into the vector pGEM1 (Promega) at the S$\text{al}$ site. Mutants of the 5’ flanking region of the yeast tRNA$^{\text{Leu}}_3$ gene were assembled from a set of 12 synthetic oligodeoxynucleotides (14),...
substituting the wild-type oligodeoxynucleotides with those carrying the desired mutation.

Preparation of extract

Yeast extract containing TFIIIB, TFIIIC and Pol III was prepared essentially according to Fabrizio et al. (13) from proteinase-deficient Saccharomyces cerevisiae strain PBJ926 (prb1-1122 prcl-407 pep4-3 can1 gal2 his1*trp1*) (15). The cells were disrupted by agitation in a Bead-Beater for a total of 10 min with intermittent cooling. The extract, in 650 mM ammonium sulfate, was loaded onto an A25 DEAE–Sephadex column (Pharmacia) that had been previously equilibrated with buffer A [50 mM Tris–HCl, pH 7.9, 100 mM ammonium sulfate, 0.1 mM dithiothreitol (DTT), 0.01 mM phenylmethylsulfonyl fluoride (PMSF), 25% v/v glycerol]. The peak flow-through fractions were pooled and dialyzed for 3 h against three 1 l changes of buffer KB containing 40 mM HEPES–KOH, pH 7.9, 140 mM KCl, 10 mM MgCl₂, 0.2 mM EDTA, 0.2 mM DTT, 0.01 mM PMSF and 50% v/v glycerol. The yeast extract (40–60 mg/ml protein concentration) was stored in small aliquots at −70°C and thawed immediately before use.

Transcription assay

The standard transcription reaction mixture contained 0.5 μl yeast extract and 300 ng tRNA gene-containing plasmid DNA in a final volume of 20 μl 20 mM HEPES–KOH, pH 7.9, 5 mM MgCl₂, 130 mM KCl, 0.1 mM DTT, 0.005 mM PMSF, 500 μM each of ATP, CTP, UTP and 50 μM α-[32P]GTP (5 μCi; Amersham). Reaction mixtures were incubated at 25°C for 20 min without rNTPs, to ensure maximal binding of transcription factors, and then for only 1 min more in the presence of rNTPs, so that tRNA processing activities were kept at a minimum. Reactions were terminated by addition of SDS to 0.5% and protease K to 0.1 mg/ml. After incubation for 30 min at room temperature and the addition of 5 μg S. cerevisiae yeast tRNA, the samples were phenol extracted and precipitated with 3 vol ethanol at −70°C for 1 h. The RNA products were subjected to electrophoresis in 8–10% denaturing polyacrylamide gels. The relative amounts of transcripts were quantified by excising the RNA bands from the gel and measuring Cerenkov radiation or by using the PhosphorImager (Molecular Dynamics) Image Quant V3.3 software. Reported transcription efficiencies are based on five independent experiments, normalized to the wild-type.

5'-End RNA mapping by primer extension

Primary in vitro tRNA transcripts were purified on 10% denaturing polyacrylamide gels, eluted in water, filtered and ethanol precipitated. Oligodeoxynucleotides (complementary to positions 1–31 or 32–53 of the tRNA^{Leu}₁ gene sequence) were 5'-end-labeled with [γ-32P]ATP. Labeled primer (20 000 c.p.m.) was added to the tRNA primary transcripts (from three standard transcription reactions) in a mixture containing 500 μM each dNTP, 10 mM DTT, 50 mM Tris–HCl, pH 8.3, 75 mM KCl and 3 mM MgCl₂. After 2 min at 37°C 100 U Moloney murine leukemia virus reverse transcriptase ( Gibco-BRL) was added to the mixture. Incubation was for 10 min at 60°C and the reaction was stopped by 5 min at 95°C. The primer extension products were phenol extracted, ethanol precipitated and fractionated by electrophoresis on 6% sequencing gels. Sequencing reactions using the same oligodeoxynucleotides were used as markers.

RESULTS

The conserved sequence element (CSE) is not sufficient for start site selection

We analyzed a large collection of S. cerevisiae tRNA genes to derive a conserved sequence element (CSE) in the neighborhood of the transcriptional start site (Fig. 1A), as others have done previously (16–17). The sequence at the start site of the tRNA^{Leu}₁ gene conforms well with the CSE; the matching block TTTCAAC is located at positions −15 and −9 upstream of the mature coding sequence and includes the start site (the purine A at position −11) indicated by the arrow in Figure 1B. To investigate whether the CSE might play a role in selection of the proper initiation site of the tRNA^{Leu}₁ gene by Pol III we constructed 5’ flanking sequence mutants translating the CSE either upstream or downstream of its original position and tested their activity using an in vitro transcription assay which had been optimized on the wild-type tRNA^{Leu}₁ gene. Transcription efficiencies of mutants were normalized relative to the wild-type. Shifting the TTTCAAC sequence 4 bases upstream increases the distance between the CSE and the A box promoter element. The transcriptional efficiency of this spacing mutant (cse-4) was 60% of the wild-type template (Fig. 2A, lane 2). However, the transcriptional start site was no longer within the CSE; instead, the purine G at position −11 in this mutant (the original start site position) was used. When the CSE sequence was moved 4 bases downstream, to decrease its distance from the mature tRNA coding sequence, the resulting mutant (cse+4) showed a strongly reduced transcriptional efficiency (16% of the wild-type) and the start site is shifted to a new position, 7 nt upstream of the mature tRNA coding sequence (Fig. 2A, lane 3), still using the same A residue within the CSE that is used in the wild-type. A primer extension analysis was performed to verify the new initiation position in the cse+4 mutant (Fig. 2B). Positioning the CSE further downstream (cse+8) reduced transcriptional efficiency to 62%; moreover, the purine at position −11 in this mutant was selected as the start site (Fig. 2A, lane 4).

These results (Fig. 2C) indicate that the presence of the conserved sequence does not ensure that transcription will start within the CSE itself. In fact, in the cse+8 and cse−4 mutants the start site position (−11) was the same as in the wild-type (Fig. 2A, lane 1). Only for the cse−4 mutant was the transcriptional start site position changed with respect to the A box internal promoter element, falling at the original A within the CSE, but then transcription efficiency was very low. The transcription apparatus appears to enforce a particular spacing between the start site and the A box, with limited flexibility (see Discussion).

Importance of a pyrimidine-purine sequence at the start site

It is known that prokaryotic and other eukaryotic RNA polymerases generally use a purine as the starting nucleotide of transcripts (18–21) and there are indications that this may also be true for Pol III (22,23). In all of our CSE spacing mutants the initiation site nucleotide was a purine preceded by a pyrimidine. To test whether a PyPu sequence is required for start site selection, we altered the PyPu motif at positions −12/−11.

We created a point mutant in which the purine A at the original start site (position −11) was replaced by the pyrimidine T. The result was a shift of the start site from position −11 to position −10, where a PyPu motif was regenerated (Fig. 3B and C).
transcriptional efficiency of this mutant was 60% of that of the wild-type (Fig. 3A, lane 5). On replacing the pyrimidine C at position -12 in the wild-type sequence with a purine A, the start site was shifted upstream to position -12, where again a PyPu motif is regenerated (Fig. 3B), and the transcriptional efficiency was reduced to 50% (Fig. 3A, lane 2). In a transition mutant in which the pyrimidine C at position -12 was replaced with the pyrimidine T, both start site and transcription efficiency were like that of the wild-type (Fig. 3A, lane 3). These results establish the importance of the PyPu motif for selection of the start site.

Combining the results presented in Figures 2 and 3, we note that very limited shifting of the position of the PyPu motif is tolerated, although transcriptional efficiency may be considerably affected. The distance separating the PyPu motif from the A box is important for good transcriptional efficiency at the correct start site. In the cse+4 mutant (Fig. 2C) the PyPu motif is selected even though its position is 4 nt downstream of that of the wild-type. To verify that the strong reduction in transcriptional efficiency (to 16% of the wild-type) of the cse+4 mutant was due to the absence of a PyPu motif at position -11, we constructed a double mutant in which a new PyPu motif was regenerated at the original position by substituting the pyrimidine T at -11 with a purine G (Fig. 2C). The start site in this double mutant (cse+4/G) was shifted back to the original position (-11) and transcription efficiency was increased to 58% of the wild-type (Fig. 2A, lane 5).

From these results we conclude that the presence of a PyPu motif at a relatively fixed distance (−11 for tRNA^{Leu}) from the first nucleotide of the mature tRNA coding region is required for high transcriptional efficiency.

**Importance of A/T base pairs near the start site**

A TTT triplet is present in the CSE just upstream of the initiation site (PyPu motif) (Fig. 1A). We investigated its effect on transcriptional efficiency and selection of the start site. We constructed a mutant (F4) in which the TTT sequence at positions −15, −14, −13 was converted to an AAA sequence (Fig. 4B). This conversion did not significantly affect the efficiency of transcription (Fig. 4A, lane 6). Replacing the TTT triplet instead with a GGC sequence drastically reduced the transcriptional efficiency of the resulting mutant (F1), to 15% of the wild-type (Fig. 4, lane 2). These results indicate that T:A or A:T base pairs upstream of the PyPu initiator region can improve transcriptional efficiency.

We then asked what the effects of the positioning of the small T:A block might be. Two double mutants were constructed based on the F1 mutant, in which TTT triplets were created in different positions upstream of the start site. In the first mutant the substitution of 2 bases (CG) at positions −18 and −17 with two Ts generated a TTT triplet 4 bases upstream of its position in the wild-type 5′ flanking region (F2 mutant). In the second double mutant substitution of the bases GGC at positions −23, −22, −21 with three Ts generated a TTT triplet 8 bases upstream of the triplet's original position in the wild-type (see Fig. 4B). The F2 mutant manifests -60% of the wild-type transcriptional efficiency, while the transcriptional efficiency of the F3 mutant was
still only 20% of the wild-type (Fig. 4, lanes 3 and 4). In all these
mutants the initiation site is the same as in the wild-type. These
data suggest both the presence and position of three T:A base pairs
can affect transcriptional efficiency, but not start site selection.

The 5' flanking sequence in *S. cerevisiae* tRNA genes are
generally TA-rich. In the yeast trnA*Leu* gene there is an AAA
triplet just downstream of the start site at positions -8, -7 and -6
(Fig. 1A). We constructed a mutant (F5) substituting the AAA
triplet with a CGG sequence (Fig. 4B). The mutation had
negligible effect on transcriptional efficiency (95%) (Fig. 4, lane
7) and did not change the initiation site.

The results indicate that T:A base pairs are important for
efficient *in vitro* transcription of tRNA*Leu*, but with limited
positioning possibilities upstream and none downstream of the
initiation site.

**DISCUSSION**

Previous studies have shown that sequence substitutions around the
transcriptional start site can have strong effects on Pol III
transcription. There is an element around the transcriptional start
site of *S.cerevisiae* 5S genes known as the 'start site element', or
SSE, which is just as significant for promoter strength as box C

(24). The influence of sequences around the start site on
transcription of genes containing internal A and B boxes has been
noted in the adenovirus VA1 gene (25), in the *S.cerevisiae* tRNA*Leu*
gene (26) and in a *Dictyostelium* tRNA*Val* gene
introduced into *S.cerevisiae* (27). The A box internal promoter alone
is not sufficient to define the initiation site accurately.

Mutations in the vicinity of the start sites of yeast tRNA genes can
have strong effects on transcription efficiency (rates) *in vitro*. For
all yeast tRNA genes where the transcriptional start site has been
mapped, the distance to the A box has been seen to be relatively
constant (between 16 and 19 bp), whereas its distance to the B box
is more variable. These observations prompted us to search for
another element(s) which might work in concert with the A box to
initiate tRNA transcription. There is conserved sequence in the
vicinity of the start site which is shared by a large number of
*S.cerevisiae* tRNA genes; 60 out of 115 tRNA genes have this CSE
around the putative start site. Our results show that the CSE acts to
select the correct initiation site only if it is positioned within a
narrow range of distances from the A box element. However,
shifting of the start site away from its normal -11 position is
accompanied by reduced transcription. When the CSE is moved
outside the specified window, the transcriptional apparatus responds
by searching for an adequate substitute within the window. It is
known that fixing the position of TFIIIC at box A is crucial for
selecting the location of the transcriptional start site; the 135 kDa
subunit of TFIIIC brings TFIIIB into the transcription complex and positions it upstream of the transcriptional start site (28). In all the spacer mutating putative purine (the start site) was always preceded by a pyrimidine and transcription experiments with transversion mutants have shown (Fig. 3A) that the PyPu motif is required for start site selection. These results also suggest that the spacing between the A box and the initiation site is relatively flexible (29).

In prokaryotes the E.coli RNA polymerase core might also contribute to promoter recognition by base pair-specific protein-DNA contacts (30), so there exists the possibility that Pol III itself might contribute some specific DNA sequence preference in the vicinity of the transcriptional start site. Even when Pol III is uniquely positioned it can scan within a small DNA segment for a preferred initiation site.

We conclude from these experiments that for the yeast tRNA genes, there are five TA-rich (75 out of the 115 genes surveyed), but the data for yeast do not indicate any clear correlation between TA content and transcription efficiency. It has recently been shown that TA-rich sequences in a variety of silkworm class III templates direct transcription at a high level (31).

Our results demonstrate that there exists a positional requirement for a TTT/AAA triplet in the 5' flanking region, which has to be both upstream and not too far from the start site for good transcriptional efficiency. This positional requirement does not influence start site selection. It is interesting to note that the substitution of a single T with a G or C at position –13 in the TTT triplet reduces transcriptional efficiency to 75% of the wild-type, without changing the start site (data not shown). The upstream versus downstream positional effect of substituting the T:A block with a G:C block seems to indicate that sequences upstream of the start site of tRNA genes may be directly correlated with promoter open complex formation. It has been shown that in tRNA transcription promoter opening and transcription bubble formation are temperature-dependent and readily reversible processes, in which 20 bp of DNA become sufficiently unwound (12). The upstream portion of the transcription bubble, in fact, appears to melt at a lower temperature than the downstream one. Preliminary experiments to determine the rates of ‘open’ complex formation in the different TTT triplet mutants indicate that active transcriptional complexes (open complexes) do not form readily, if at all, in the F1 and F3 mutants, whereas they do in the transcriptionally efficient wild-type and F2 mutant. Results of a template competition experiment using a reconstituted system showed that transcription from either the F1 or the F3 mutants was not reduced by the addition of a wild-type gene, which means that both TFIIIC and TFIIIB are stably associated in both mutant genes (data not shown).

We conclude that selection of the start site and efficiency of transcription in tRNA genes in vitro are directed by an initiation window established by TFIIIB, which directs the catalytic subunits of Pol III toward the nucleotides comprising the window itself (32). The position of this initiation window is slightly different for tRNA genes with different 5’ flanking sequences. One initiator sequence might be WWN_p-2-PyPu, (which is present in 104 out of 115 genes surveyed on the EMBL database). It is interesting to note that 104 out of 115 5’ flanking sequences of tRNA genes have the PyPu motif at a position between –12 and –10. In S.cerevisiae the transcriptional start sites (5’-end RNA mapping) are known for only 10 tRNA genes and their position is between –12 and –10 nt upstream of the first nucleotide of the mature tRNA coding sequence. The role of this ‘initiator’ region will be understood by studying, in an vitro system and using highly purified components, the interaction between these nucleotides and the catalytic subunit of Pol III.

ACKNOWLEDGEMENTS

We thank E.P. Geiduschek and members of his laboratory for critical reading of the manuscript and K. Williams for helpful discussions and for help with the manuscript. We also thank G. Di Franco for technical assistance and A. Sebastiano for secretarial assistance. This research was supported by Progetto Finalizzato Ingegneria Genetica. G.P. was supported by a fellowship from the Fondazione Adriano Buzzati-Traverso.

REFERENCES