Nucleic Acids Research, Vol. 19, No. 19 5331–5337

# Dual modes of transcriptional and translational initiation of *SSP1*, the gene for a mitochondrial HSP70, responding to heat-shock in *Schizosaccharomyces pombe*

Hiroaki Kasai[1] and Katsumi Isono[1,2]*
[1]Postgraduate School of Science and Technology and [2]Department of Biology, Faculty of Science, Kobe University, Rokkodai, Kobe 657, Japan

Received June 20, 1991; Revised and Accepted September 3, 1991 EMBL accession no. X59987



## ABSTRACT

**The *SSP1* gene of *Schizosaccharomyces pombe* which is homologous to the *SSC1* gene of *Saccharomyces cerevisiae* was cloned and its nucleotide sequence determined. A heat-shock element and three possible TATA boxes were found upstream of the coding region. Dual modes of transcriptional initiation were observed in primer extension analyses using as templates the mRNAs prepared from cells before and after heat-shock. Initiation sites situated 50 to 60 nucleotides downstream of the normal one were found to be additionally used in cells after heat-shock. Thus, the mode of transcription in heat-shocked cells of *S. pombe* appears to be more similar to the one observed in higher eukaryotes than that in *S. cerevisiae*. The *SSP1* gene contains two methionine codons in the region predicted to encode a mitochondrial targeting signal sequence of its protein (Ssp1) and the stretch between the two methionine codons is capable of forming stem-loop structures. Both of the two methionine codons were confirmed to function as translational initiators *in vitro*. A possible mechanism is proposed for the response of the *SSP1* gene towards heat-shock by the differential initiation of its transcription and translation.**

## INTRODUCTION

The genes for heat-shock proteins comprises a multi-gene family in various organisms and are differently expressed under a variety of physiological conditions. Most notably, when cells are exposed to a higher temperature, they become highly expressed. The products of these genes have been established to interact with many proteins and protein complexes such as the DNA replication complex (1, 2), ribosomal proteins (3), premature IgG heavy chains (4), and a subunit of endonuclease (5). Nine genes have so far been classified into one of the heat-shock protein families, the *HSP70* gene family, in *Saccharomyces cerevisiae* (6). The *SSC1* gene is one of them. Its product, Ssc1 protein, is known to be localized in mitochondria (7). By genetic analyses Ssc1 protein was shown to be required for the import of several precursor proteins into mitochondria (8). It seems to be a key initial component involved in the chain of reactions leading to refolding of proteins transported into mitochondria.

It should be noted, however, that Ssc1 protein was detected also in the nucleus after heat-shock (5). Generally, it has not yet been clarified how proteins are transported into the nucleus, although some sequence motifs have been proposed to be required for nuclear localization of proteins. Hall et al. (9), for example, proposed a nuclear targeting sequence containing two positively charged amino acid residues flanking three hydrophobic residues including a proline. This sequence was found to be present in several nuclear proteins of *S. cerevisiae* such as α2 protein (10), histone H2A (11), histone H2B (12), and histone H4 (13). From the amino acid sequence analysis of Ssc1 (5), a sequence, RLIPR, satisfying the criteria for this putative nuclear targeting signal was found. However, it remains to be investigated further why this protein was detected in the nucleus after heat-shock and what a function this protein plays there. It might be that Ssc1 protein plays a role in the nuclear transport of some proteins which are required in the nucleus after heat-shock.

Our laboratory has been performing cloning and characterization of the mitochondrial ribosomal proteins encoded by the nuclear genes in *S. cerevisiae* (14–17) and *S. pombe* (Kasai and Isono, unpublished results) and investigated their structural characteristics, transport into mitochondria, evolutionary conservation, and so on. In spite of some notable differences between the two yeasts, structural and functional similarities in the mitochondrial ribosomal proteins, mitochondrial RNA polymerase and mitochondrial heat-shock proteins have been observed (Kasai and Isono, unpublished results). To facilitate further our studies on the mechanism regulating the expression of the nuclear genes for mitochondrial ribosomal proteins and other proteins associated with the mitochondrial ribosomes in *S. cerevisiae*, we raised antibodies against a mixture of proteins of the mitochondrial ribosomes prepared from commercial baker's yeast cells. By using the resultant antibodies in screening cDNA clones, we obtained a clone which contained the gene for a mitochondrial Hsp70. Upon nucleotide sequencing of this clone, we found that the gene was identical to the *SSC1*

* To whom correspondence should be addressed

gene reported by Craig et al. (7). Since Morishima et al. (5) found the presence of a possible nuclear-targeting signal in the protein (Ssc1) encoded by this gene, we were interested to investigate its intracellular localization and possible roles in different cellular compartments. For this purpose, we performed comparative studies with a homologous protein from *S. pombe* as will be described below.

## MATERIALS AND METHODS

### Strains, culture media and antibody preparation

The *S. pombe* strain 972 ($h^-$) was obtained from M.Yanagida of Kyoto University. For the isolation of *SSC1* cDNA clones, total poly(A) RNA was prepared from cells of *S. cerevisiae* strain 07173 (**a**/$\alpha$) and a cDNA library was constructed. Commercial baker's yeast cells were kindly provided by Oriental Yeast Company (Osaka) and mitochondrial ribosomal proteins were prepared from them. The resultant ribosomal proteins were mixed with Freund's adjuvants (Difco) and used to raise antibodies against them in a male rabbit. When necessary, antibodies specific to Ssc1 protein were purified by the method of Smith and Fischer (18). YPD medium [1% Bacto-yeast extract (Difco), 2% Bacto-peptone (Difco), 2% dextrose] or YPG medium [1% Bacto-yeast extract (Difco), 2% Bacto-peptone (Difco), 3% glycerol] was used for cultivation of yeast cells.

### Preparation of DNA and RNA

The genomic DNA was prepared as described by Niederacher and Entian (19). Total RNA was prepared essentially as described by Sprague et al. (20). After extraction, nucleic acids were concentrated by ethanol precipitation and total RNA was purified by lithium chloride precipitation.

### Primer extension analysis

Two oligonucleotide primers, 5'-CCGATTACGGGACCCTT-AAC-3' (primer #1) and 5'-CCAGAAAAGCACAATAGAA-ACC-3' (primer #2), were synthesized using a DNA synthesizer from ABI (model 381A) and their 5'-termini were labeled with [$\gamma$-$^{32}$P]ATP. 120 $\mu$g of total RNA and one of the end-labeled primers ($1\times10^6$ cpm) were mixed, denatured at 85°C for 10 min and then incubated at 30°C for 12−14 hrs in 30 $\mu$l of 40 mM PIPES (pH 6.54) containing 400 mM NaCl, 1 mM EDTA and 80% formamide. The primer-RNA hybrid was precipitated with ethanol and redissolved in 24 $\mu$l of 50 mM Tris-HCl (pH 8.3 at 42°C) containing 10 mM $MgCl_2$, 10 mM dithiothreitol, 60 mM NaCl, 2.5 mM each of four dNTPs and 10 units of AMV reverse transcriptase (BRL). Primer extension reaction was carried out at 42°C for 1 hr. After degrading RNA by the incubation at 75°C for 20 min with 150 $\mu$l of 0.15 M NaOH/5 mM EDTA, the primer-extended DNA products were run on a sequencing gel alongside a dideoxy sequencing ladder prepared with the same primer.

### *In vitro* transcription and translation

Transcription with T3 RNA polymerase of pSSP1Δ1 which had been linearized by digestion with *Pvu*II and *Hinc*II was performed as described (21). The RNA transcripts were processed with an mRNA capping kit (Stratagene). A 2 $\mu$g portion of the synthesized RNA was added to 30 $\mu$l of a *in vitro* translation mixture containing the rabbit reticulocyte translation kit (BRL) and 10 $\mu$Ci of Tran$^{35}$S-LABEL™ (1000 Ci/mmol: ICN). Its salt concentrations were adjusted to 73 mM $CH_3COOK$ and 0.87 mM Mg-acetate. After incubation at 30°C for 60 min, an equal volume of 2×sample loading buffer [100 mM Tris-HCl (pH 6.8) containing 200 mM dithiothreitol, 8% SDS, 0.2% bromphenol blue and 20% glycerol] was added and the mixture was boiled for 5 min. The reaction mixture was cleared by centrifugation at 15000 rpm for 10 min and the synthesized products were subjected to electrophoresis in a 12.5% polyacrylamide gel in the presence of 0.1% SDS. The gel was soaked for 1 hr in a one- to three-fold volume of EN$^3$HANCE (NEN) and the protein bands were visualized by fluorography.

## RESULTS

### Cloning of *SSP1* and analysis of its 5'-untranslated region

A fragment of the cDNA clone containing the *SSC1* gene of *S. cerevisiae* which we cloned as described above was found to cross-hybridize with the genomic DNA of *S. pombe* (Fig. 1). Therefore, we isolated a plasmid clone, termed pSP22, which contained the desired *S. pombe* gene. To characterize the gene further and to study its expression under various conditions, we tried to isolate a genomic DNA clone containing the gene and its 5'-untranslated region. Upon double digestion with *Bam*HI and *Eco*RI, a genomic DNA fragment of 1.95 kb in length was found to hybridize with the 5'-end of the *SSC1* gene. Genomic fragments of this size were therefore cloned into pUC119 and named pSP22up (Fig. 2). While we were performing these experiments, Powell and Watts (22) reported the cloning and nucleotide sequencing of an *S. pombe* gene which they named *SSP1*. This gene appeared to be identical to the one we had isolated. Indeed, our nucleotide sequence data (not shown) of the gene in plasmid pSP22 perfectly matched their data for the

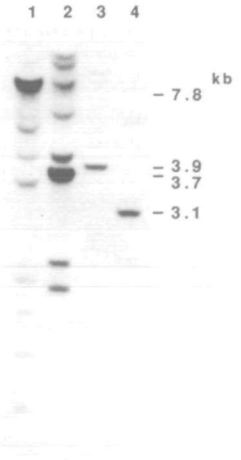


**Figure 1.** Southern hybridization analysis of the nuclear genes encoding mitochondrial Hsp70 in *S. cerevisiae* and *S. pombe*. Ten micrograms of the *S. cerevisiae* genomic DNA (lanes 1 and 2) or the *S. pombe* genomic DNA (lanes 3 and 4) were digested with *Eco*RI (lanes 1 and 3) or *Hin*dIII (lanes 2 and 4) and subjected to electrophoresis in a 0.8% agarose gel. Hybridization was performed overnight at 37°C in a solution containing 25% formamide, 5×SSC, 1×Denhart's solution, salmon sperm DNA (100 $\mu$g/ml), 0.1% SDS and a $^{32}$P-labeled *SSC1* probe. The membrane was washed at 50°C four times, each time for 30 min, with 5×SSC containing 0.1% SDS. The positions of molecular size markers (in kilobases) are indicated at right.

coding region of *SSP1*. Therefore, we use hereafter the same nomenclature to indicate this gene.

The nucleotide sequence of clone pSP22up was determined after the introduction of unidirectional deletions. The sequence data (Fig. 3) indicate that the clone indeed contains the 5′-untranslated region of *SSP1*. We found that the cloned segment in pSP22up contained an additional open reading frame which was situated at positions −539 through −736 on the opposite strand as shown in Fig. 3. This open reading frame did not show a significant degree of similarity to any known genes in both the GenBank and the NBRF databases. Our nucleotide sequence data of the 5′-untranslated region differed from those of Powell and Watts (22) at two positions: their data suggest the presence of only one T at position −94, whereas our data suggest the presence of two T's at this position. Conversely, the sequence of Powell and Watts (22) at position −128 through −131 is TTTTT, whereas our data suggest that it is TTTT.

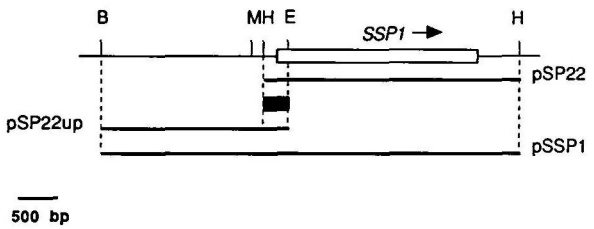


**Figure 2.** Restriction map of the upstream region of *SSP1*. Open bar indicates the coding region of *SSP1* and arrow shows the direction of its transcription. Cross hatched bar shows a probe used for the isolation of clone pSSP1, in which pSP22 and pSP22up were linked. Restriction enzyme sites are *Bam*HI (B), *Mlu*I (M), *Hin*dIII (H), and *Eco*RI (E).

```
-1051 - -1021                                  GTGACTTTTGGATGAAGCAAAGGTGAGCAAT
                                               :  :  :  :  :  :  :  :  :  :  :
-1020 -  -961  TAAATTTCACAGATGAATTACATTTAGGAGAAAAGATCAGTATCCAAAAATATAATATTA
                                               :  :  :  :  :  :  :  :  :  :  :
 -960 -  -901  CTAGCTATTTTTTTTTTCTCAAACTTGGCAGTTCCTTTAATTCTCTCTTGATAATAAAAA
                                               :  :  :  :  :  :  :  :  :  :  :
 -900 -  -841  TATTAGAAATTCTACTACCCACTGTATATTTTATTCCTTTCGGAATGTTCTTCAGATTGA
                                               :  :  :  :  :  :  :  :  :  :  :
 -840 -  -781  CCCAAGAAATATAGAATAAGCTAAAATTAACAGATTATACTTTATTTTCTTTAATAATTG
                                               :  :  :  :  :  :  :  :  :  :  :
 -780 -  -721  TACTAATTTCGGTTTCTTAATTGCTACTTAATCCACTTAGTGGATTATAGGAATTTCCAT
                                                                 *  L  F  K  W
 -720 -  -661  AAATTTTCCACTGCCAATCTTGAATACATTTTCCTTTTTTTCTCTTTCATTTGAATGCTT
               L  N  E  V  A  L  R  S  Y  M  K  R  K  K  E  K  M  Q  I  S
 -660 -  -601  CTATCCCTTTTCAGTAAATACTGAGGTGGCATTTTTCTATATCTTGCCAATAGTCCAATT
               R  D  R  K  L  L  Y  Q  P  P  M  K  R  Y  R  A  L  L  G  I
 -600 -  -541  CTTTTCCATTTGTTTATCAAATACTGAATTATAACGAGAGCAACATTAAGCATTTTTATC
               R  K  W  K  N  I  L  Y  Q  I  I  V  L  A  V  N  L  M  K  I
 -540 -  -481  ATGTAACCACAGGCTTCTTAAAACGGTTTAGTACCGTATTACTGCAAACCTTCTTCGAGC
               M
                                               :  :  :  :  :  :  :  :  :  :  :
 -480 -  -421  TTCCTTTTCCTTCACCATCGCACCATCCCGCTTGTCCATTTTTCTCTGGCATTCGATGTT
                                               :  :  :  :  :  :  :  :  :  :  :
 -420 -  -361  AACCGAACAGTAAGGGGAGGGCACTAGTTAGTCCACTTAATTTTTTTGATGAGGTAGCGA
                                               :  :  :  :  :  :  :  :  :  :  :
 -360 -  -301  TTGCCACGAAGAAGGTTCTCGTTGTTAAGGTACTTGTTTTTTTGTAACCAGAACCACTAC
                                               :  :  :  :  :  :  :  :  :  :  :
 -300 -  -241  CTCCCTTGGGAATTTAAAAAGAGGAATACGCGTCGTTTTTGTGGAAAAGCGCACATTAATA
                                               :  :  :  :  :  :  :  :  :  :  :
 -240 -  -181  TATCCCAATAGTTAGAGAAAAGCAAAGTTTTTCTGGTTTCTATTGTGCTTTTCTGGAAAA
                                               :  :  :  :  :  :  :  :  :  :  :
 -180 -  -121  CATATATTTTCTTCGCAACTCTTTTATTAAAGCTTTTAGGAATTTTGAATTTTGATCGAA
                                               :  :  :  :  :  :  :  :  :  :  :
 -120 -   -61  TTTTAGAAAAAACTATTCGCAAGACTTACAATTTTTGAAGGGTGCTATTTGTGAAAAAAT
                                               :  :  :  :  :  :  :  :  :  :  :
  -60 -    -1  AAAACGTGAAATAAATCGTTTTATAATTTACGAATTGTCGTTATTCAAAACTCAAAAAAT
                                               :  :  :  :  :  :  :  :  :  :  :
    1 -    60  ATGATCTCGTCGAGATTCACTAATGTAGTCCGTAGCGGATTGCGTTTCCAAAGCAAGGGA
               M  I  S  S  R  F  T  N  V  V  R  S  G  L  R  F  Q  S  K  G
                                               :  :  :  :  :  :  :  :  :  :  :
   61 -   120  GCATCGTTCAAGATTGGCGCTTCCTTGCATGGAAGTCGCATGACCGCCCGCTGGAATTCT
               A  S  F  K  I  G  A  S  L  H  G  S  R  M  T  A  R  W  N  S
```

**Figure 3.** Nucleotide sequence of the 5′ region of *SSP1* gene. The nucleotide sequence of clone pSP22up was determined. Nucleotides are numbered starting at the A of the first ATG initiator codon (see text) and the first 40 codons are translated. The nucleotides before the ATG are negatively numbered. The nucleotides marked with black and gray arrow-heads indicate, respectively, the transcriptional initiation sites in cells with or without heat-shock (see Fig. 4). A putative heat-shock element (HSE) is indicated by black bar and a putative cAMP responsive transcriptional control element (23) is indicated by gray bar. Likely TATA boxes are enclosed. The amino acid translation of an additional open reading frame found in this region in the reverse direction is shown above the nucleotide sequence.

## Analysis of the transcript of *SSP1*

Powell and Watts (22) reported the presence of a TATA box-like sequence at positions corresponding to −36 to −39 of our sequence data in Fig. 3. They also predicted the presence of three heat-shock element-like sequences in the −50 to −150 region. To investigate whether these putative *cis*-elements would actually be functional, primer extension analysis was performed with four RNA templates, namely the RNAs prepared from cells with or

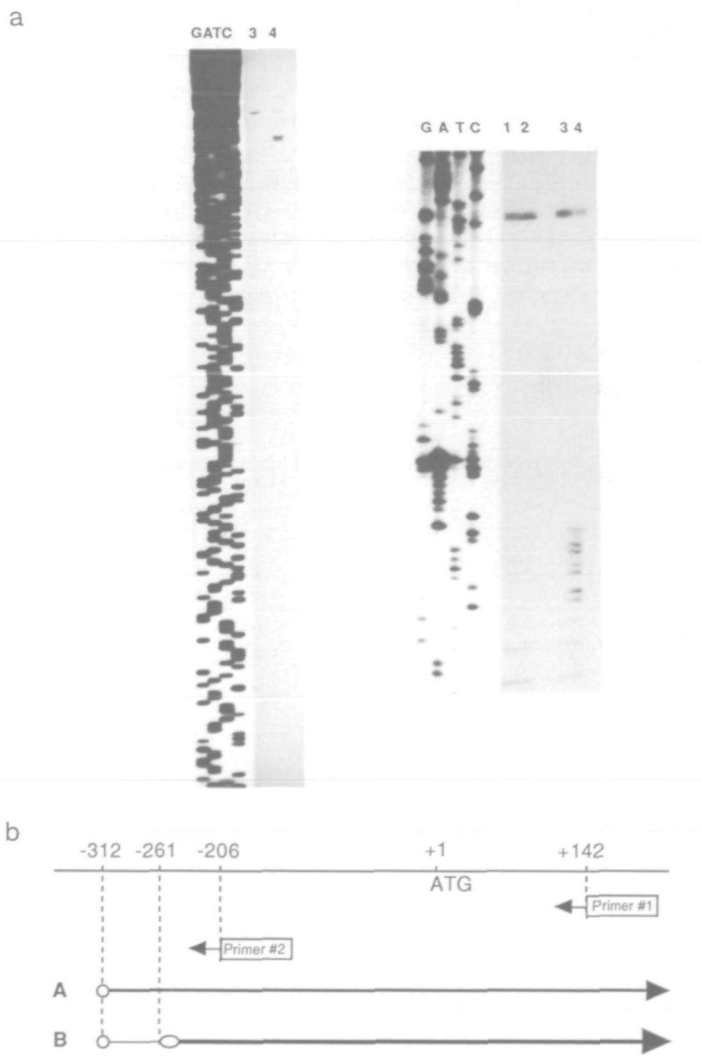


**Figure 4.** Primer extension analysis with the *SSP1* transcripts under different growth conditions. a. Primer extension analyses were performed using either primer # 1 (left panel) or primer # 2 (right panel). The primer-extended DNA products were subjected to electrophoresis in a sequence gel in parallel with a dideoxy sequencing ladder produced with the corresponding primer. As templates 120 μg of total RNA from cells grown continuously at 24°C (lane 3) or from cells shifted to 37°C and cultivated for 30 min (lane 4) were used. In lanes 1 and 2 (right panel) total RNA from cells grown at 28°C in YPD and YPG, respectively, were used. b. The results of the primer extension analyses are schematically illustrated. Open circles represent the transcriptional initiation sites of *SSP1* in cells grown without heat-shock (A) or with heat-shock (B).

without heat-shock, and the RNAs from cells cultured with or without glucose. A synthetic oligonucleotide corresponding to nucleotides at positions +142 through +161 was used as a primer. It was found that *SSP1* transcripts were initiated at a position unexpectedly far away from the primer-binding site. Moreover, the major *SSP1* transcripts present in cells after heat-shock were initiated at sites distinctly different from those under other three growth conditions (Fig. 4a).

To determine the transcriptional initiation sites more precisely, an oligonucleotide primer binding to a stretch from −206 through −185 was synthesized and the primer extension analysis repeated. Consequently, the *SSP1* transcript was shown to start at a C positioned at −312 when cells were not exposed to heat-shock, regardless of whether they were cultured with or without glucose. However, when cells were heat-shocked, in addition to the C at −312 several positions between −261 to −251 became major initiation sites for the *SSP1* transcription (Fig. 4b). Because we applied an equal amount of RNA to each lane and because the total RNA preparations used in these experiments contained large amounts of ribosomal RNA, the total amount of *SSP1*-specific transcripts in lane 3 (without heat-shock) and that in lane 4 (with heat-shock) can directly be comparable. The latter appears to be about twice as much as the former. The nucleotide sequences surrounding the two regions in which transcriptional initiation occurs are similar to each other (Fig. 3). Therefore, these data were interpreted to indicate that switching of transcriptional initiation site occurs in *S. pombe* cells when they are heat-shocked which results in a quantitative and qualitative difference in the *SSP1* transcription.

## The Mitochondrial Targeting Sequence of *SSP1*

The coding sequence of *SSP1* is likely to start at an ATG codon situated at the position indicated as +1 in Fig. 3 as this is the first ATG codon of this open reading frame. It was noticed that the G+C content of the nucleotide sequence of *SSP1* changes drastically from 25% to 46% after this ATG codon. There is an additional in-frame ATG codon 96 nucleotides downstream of the first ATG codon. The latter is the one proposed to be the initiator codon of *SSP1* by Powell and Watts (22). From our primer extension analysis, a single transcript was found to be synthesized from *SSP1* when cells were not heat-shocked. *This transcript includes both of the two potential initiator codons in question (Fig. 4). Interestingly, two stem-loop structures could be formed in the region of SSP1* mRNA flanked by the two initiator codons. They are located at positions +22 through +46 and at +53 through +88 as depicted in Fig. 5b. The second ATG codon is situated after these structures. Free energy of these structures was calculated to be −7.9 kcal/mol and −17.31

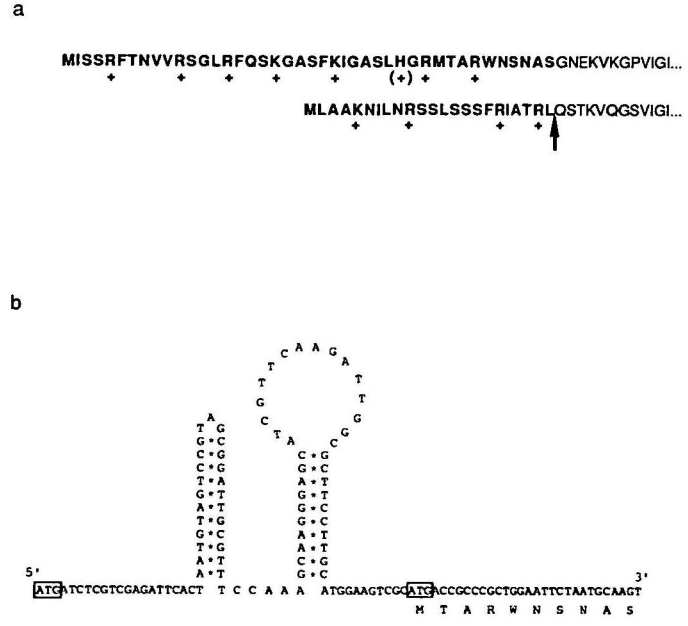


**Figure 5.** The secondary structures in *SSP1* mRNA might be responsible for the dual modes of translational initiation. a. The N-terminal amino acid sequences of Ssp1 (top) and Ssc1 proteins (bottom) are aligned. Putative mitochondrial targeting sequences are expressed in bold face. Arrow indicates the cleavage site of the mitochondrial targeting sequence of Ssc1 protein (24) and + indicates positively charged amino acid residues. b. Possible secondary structures in the *SSP1* gene coding for the putative mitochondrial targeting sequence of Ssp1 protein are illustrated. The two ATG codons are enclosed. Prediction of a most stable secondary structure and calculation of its free energy were performed by the SECST option (25) of the GENETYX program (Software Development Co., Ltd.). Nucleotide sequence positions from +80 through +260 shown in Fig. 3 were analyzed.

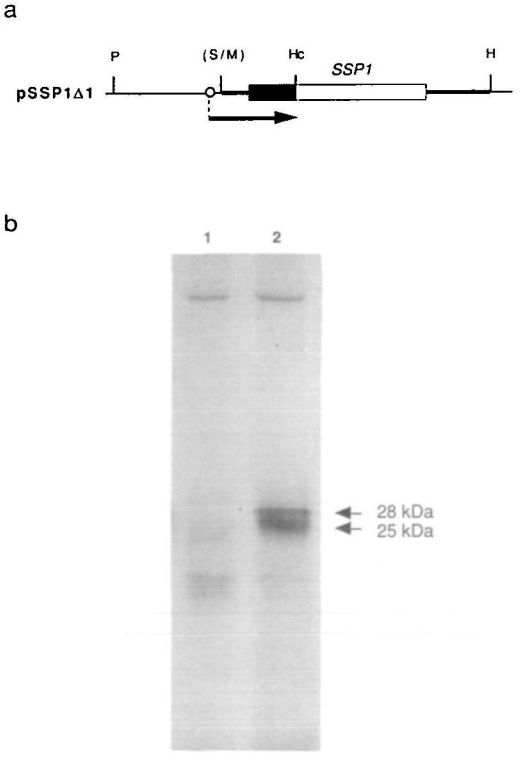


**Figure 6.** The second ATG codon in *SSP1* mRNA functions as a translational initiator. a. Plasmid pSSP1Δ1 was constructed by ligating pSSP1 DNA cleaved with *Sac*I on the vector and *Mlu*I on the *SSP1* fragment. It was transcribed *in vitro* from the T3 promoter (○) after truncation by cleavage with *Pvu*II (P) and *Hin*cII (Hc). Arrow indicates the expected truncated Ssp1 mRNA which is capable of encoding a truncated Ssp1 protein corresponding to the cross hatched region of *SSP1*. Other restriction enzyme sites are *Hin*dIII (H), *Mlu*I (M), and *Sac*I (S). b. Two micrograms of capped RNAs synthesized as described above were incubated in a rabbit reticulocyte lysate translation system in the presence of $^{35}$S-Met. The translational products were subjected to electrophoresis in a 12.5% polyacrylamide gel in the presence of 0.1% SDS and visualized by exposure to Fuji X-ray film HR-A for 1 week. Arrows show translational products. In lane 1, the products without added RNA were analyzed.

kcal/mol, respectively. Therefore, they can form moderately stable stem-loop structures. By aligning the amino acid sequence deduced from the nucleotide sequences of *SSP1* and its *S. cerevisiae* counterpart, *SSC1*, a putative mitochondrial targeting sequence was deduced for Ssp1 protein (Fig. 5a). The stretch nicely corresponds to the mRNA region capable of forming stem-loop structures. Therefore, we were interested to examine whether the translational initiation of the *SSP1* mRNA would occur at only one of the two ATG codons or both and whether it would differ in normally grown cells and in cells after heat-shock.

### *In Vitro* Translational Analysis of *SSP1* mRNA

To analyze the translation of the *SSP1* transcripts, we constructed plasmid pSSP1 and its derivatives, pSSP1Δ1. Plasmid pSSP1 was constructed by combining the chromosomal segments carried by plasmids pSP22 and pSP22up (see Fig. 2). Plasmid pSSP1Δ1 was derived from pSSP1 in which a segment from a *Sac*I site of the vector through an *Mlu*I site of the insert in pSSP1 (Fig. 2) was eliminated. In this way, the length between the T3 promoter on the vector and the *SSP1* gene was shortened. The *Mlu*I site in question is 12 bp upstream of the upper-most transcriptional initiation site found in heat-shocked cells. Therefore, the mRNAs synthesized from pSSP1Δ1 were expected to be identical to the native transcripts observed in heat-shocked cells except for the one initiated at position −312 (see above). The plasmid was truncated by cleavage with *Pvu*II and *Hinc*II and transcribed *in vitro* using T3 RNA polymerase as depicted in Fig. 6a. The truncation was intended to make the relative size difference of the translational products from the first and second initiators larger. To enhance the translational efficiency, the RNA synthesized was capped with $m^7G(5')ppp(5')G$. The resultant *in vitro* transcripts were found to be of expected size in denaturing gel electrophoresis analysis (data not shown). The RNAs thus prepared were translated in a rabbit reticulocyte lysate system and the translational products obtained were subjected to SDS polyacrylamide gel electrophoresis as shown in Fig. 6b. Two protein bands were detected. Their sizes were estimated to be 25 and 28 kDa, respectively. The values are slightly larger than the expected sizes of 19.9 and 23.5 kDa which are calculated from the nucleotide sequence data shown in Fig. 3. However, the differences are within an experimental error range. These results suggest, therefore, that both of the ATG codons indeed function as initiation codons. In a preliminary experiment, either one of the ATG codons was eliminated by changing their nucleotide sequences and the altered plasmids were similarly analyzed. The results (not shown) confirmed the conclusion mentioned above.

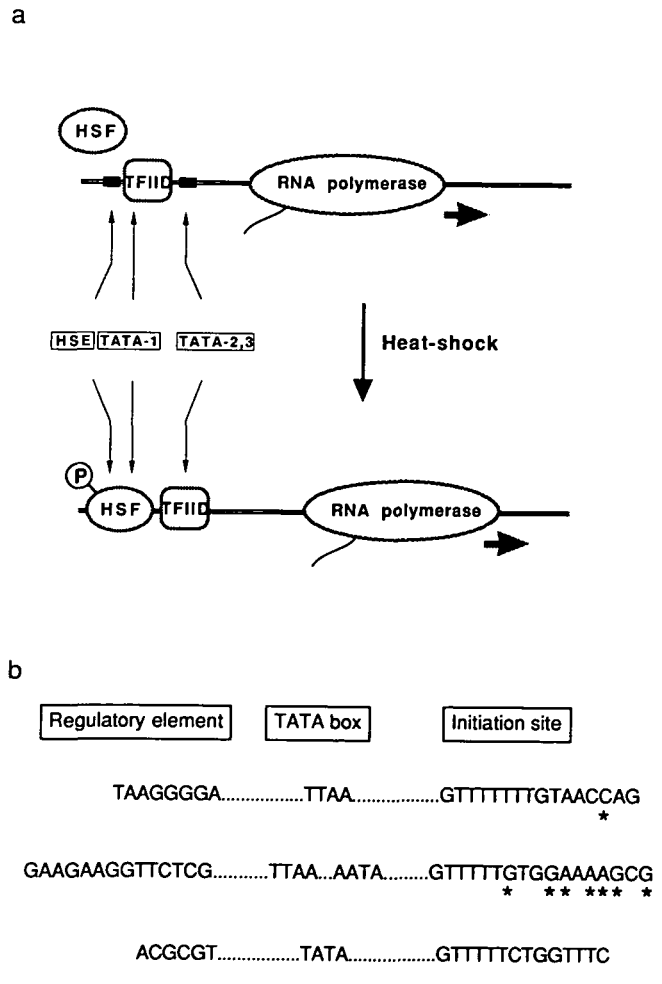


**Figure 7.** A model for transcription in response to heat-shock in *S. pombe* and sequence similarities in the transcriptional initiation sites. a. Under normal growth conditions (top) the HSE is empty and TFIID binds to TATA-1. In heat-shocked cells (bottom), however, TATA box-1 will be covered by phosphorylated (active) HSF which now binds to the HSE. Therefore, TFIID is moved to TATA box-2 or 3. As a result, transcriptional initiation site will be shifted downwards. b. Nucleotide sequences of the two initiation sites determined by the primer-extension analyses (top and middle) are aligned together with their possible regulatory elements. At the bottom a possible transcriptional initiation site with a *Mlu*I motif is shown for comparison (see the Discussion section).

## DISCUSSION

### Characteristics of the structure and function of *SSP1*

The induction of eukaryotic heat-shock genes in response to a temperature shift is known to be mediated by heat-shock elements (HSE) the sequences of which are highly conserved during evolution (26, 27). The HSE provides a binding site for the heat-shock transcription factor, HSF (28−30). In human and in *Drosophila*, HSF is activated and becomes able to bind to DNA through phosphorylation after heat-shock (31−33). In *S. cerevisiae*, on the other hand, HSF is bound to DNA in non-stressed cells as well (34) and the stimulation of transcription appears to be effected by a conformational change of the DNA-bound HSF after heat-shock (35, 36).

The 5′-untranslated region of *SSP1* was found to be much longer than those of many other genes: it is 251 to 261 bp in length in heat-shocked cells and 312 bp in normally grown cells. Also, it is low in G+C content (29.8%). Kozak (37) surveyed mRNA sequences of higher eukaryotes and found that about three-quarters of the mRNAs examined have 5′-noncoding sequences of 20 to 100 nucleotides in length and only about one-quarter of them had 5′-noncoding sequences of 100 to 300 nucleotides in length. In yeast, 5′-noncoding regions in general are low in G+C content and their average length is 52 nucleotides (38). Kozak (39) demonstrated that the preproinsulin gene placed on a simian virus 40-based plasmid showed an enhanced translation in cultured monkey cells when the 5′-noncoding region of the gene was elongated. Although it is not yet clear how the

5'-noncoding sequence affects the post transcriptional regulation, the long 5'-noncoding region might affect the translational efficiency of *SSP1* mRNA.

HSE's have been shown to exist upstream of most heat-shock responding genes in yeast and their consensus sequence has been reported to be **CNNGAANNTTCNNG** (40). In the nucleotide sequence of the 5'-untranslated region of *SSP1* (Fig. 3), an HSE-like sequence, **GAAGAAGGTTCTCG**, was found. This sequence is almost identical to the consensus sequence shown above except for the 5'-terminal G. It is present at positions −353 through −340, about 100 bp upstream of the transcriptional initiation sites identified in cells after heat-shock. Earlier, Powell and Watts (22) predicted the presence of five HSE-like sequences and a TATA-box upstream of the coding region of *SSP1*. These sites correspond to −58, −89, −120, −139 and −146 in our sequence data shown in Fig. 3. Our primer extension analysis data clearly indicate that the transcription of *SSP1* starts almost exclusively at the C located at position −312 (Fig. 4). The resultant transcript contains a translational initiator situated about 100 nucleotides upstream of the one predicted by Powell and Watts (22). The data shown in Fig. 6 indicate that both of the two translational initiators, including the one suggested by Powell and Watts (22), are functional *in vitro*.

Despite the fact that many genes in *S. pombe* analyzed so far contained a typical TATAbox as in higher organisms, such was not the case with *SSP1*. Instead, three TATA box-like sequences which are deviated from a typical TATA-box were found as indicated in Fig. 3, although it is not directly proven that they are actual transcriptional initiation sites. It is tempting to speculate that the deviation of these sequences from a typical TATA-box might play a role in the above-mentioned switching of transcriptional initiation site upon heat-shock.

It should be noted here that the nucleotide sequences surrounding the two transcriptional initiation sites are very similar to each other as shown in Fig. 7b. A stretch containing $G(T)_{5\ or\ 7}GT$ is present near both sites, i.e. at positions −267 through −260 and at positions −325 through −316. Similar sequence motifs, $A(T)_7GA$ at positions −381 through −372 and $G(T)_5C$-T at positions −214 through −207, exist in the adjacent regions. It might be that transcription is initiated at one of these motifs under different conditions. Upstream of the last motif, a typical TATA box at positions −242 through −239 and an *Mlu*I motif-like sequence, ACGCGT, at positions −274 through −269 are present. An *Mlu*I motif is known to be a *cis*-acting element which exerts a cell cycle stage-dependent transcription on several genes of *S. cerevisiae* (41). McConnell et al. (42) have isolated mutants of *S. cerevisiae* that are temperature-sensitive for the transfer of mitochondria into a growing bud. The mutations are cell-cycle specific, showing their defects in late G1 or early S phase. Therefore, it would be interesting to analyze the *SSP1* expression during the cell cycle.

Another characteristic feature of *SSP1* is the differential transcriptional initiation observed in cells after heat-shock. In heat-shocked *S. pombe* cells, the *SSP1* transcription is additionally initiated 50 to 60 bp downstream of the normal initiation site. This observation suggests that the HSF in *S. pombe* may bind to HSE only in heat-shocked cells as observed in human and *Drosophila*, because otherwise the transcriptional initiation site under normal growth conditions is blocked by HSF. In support of this notion, Gallo et al. (43) found that the *S. pombe* HSF was very similar in size and in DNA binding properties to mammalian HSFs than to the *S. cerevisiae* HSF. Thus, the transcriptional regulation of *SSP1* in response to heat-shock is postulated as illustrated in Fig. 7a. Under normal growth conditions, HSF does not bind to HSE and consequently TFIID recognizes a TTAA sequence positioned at −336 through −333 (TATA-1) and binds to it. When cells are exposed to a temperature shift, HSF is phosphorylated and becomes capable of binding to the HSE which is positioned near the putative TFIID binding site in normal cells. As a result, TFIID is predicted to move to a secondary binding site, a TTAA sequence positioned at −287 through −284 (TATA-2) or to an AATA sequence positioned at −277 through −274 (TATA-3). Experiments to examine these possibilities and to quantify the amounts of different transcriptional and translational products before and after heat-shock are currently in progress.

### Intracellular localization of Ssp1 protein

Mitochondrial Hsp70 protein has been shown to be transported into mitochondria as well as to the nucleus in *S. cerevisiae* (5). A s described in the Results section, the *SSP1* mRNA contains two AUG codons which flank two possible stem-loop structures (Fig.5). The calculated free energy values for them, −7.9 and −17.31 kcal/mol, are reasonably high. Kozak (44) constructed a plasmid containing two ATG codons and a stem and loop structure between them with a free energy of −19.4 kcal/mol and analyzed its transcript as well as protein products *in vitro*. The distance between the first ATG codon and the stem and loop structure was varied while the distance between the secondary structure and the second ATG codon was kept constant at 32 bp. When the former distance was set at 35 bp, both ATG codons were observed to function as initiation codons. These results seem to indicate that the distance between the first and second ATG codons and the stretch capable of forming a stem and loop structure are important in the translation of this gene.

In the case of *SSP1*, the putative secondary structures are located in the region 19 bp downstream of the first ATG codon and 12 bp upstream of the second ATG codon. At first, we thought that the second ATG codon is non-functional, because its −3 position was found to be occupied by a C instead of an A that was reported to be a best fit (37). However, two protein products with a difference of 3 kDa were detected as shown in Fig. 6. Moreover, the *SSP1* gene containing a point mutation in the first ATG codon was still capable of producing a protein product *in vitro* (Kasai and Isono, unpublished). These data suggest, therefore, that both ATG codons are functional. It is probable that the secondary structure of mRNA (Fig. 5) facilitates an alternative initiation of translation of the *SSP1* transcript. Some proteins are known to be partitioned between two subcellular organelles. Their genes have two transcriptional initiation sites and their messengers contain two translational initiator codons. The two translational initiators are used differently in the messengers transcribed from the different initiation sites (45−47). The Ssp1 protein is transported into mitochondria and, most probably, into the nucleus as in the case of the Ssc1 protein of *S. cerevisiae* (5). Its sorting might be dependent on whether the protein has a mitochondrial targeting signal which is created when translation starts at the first initiator codon or a shorter one when translation starts at the second initiator codon. This will in turn be influenced whether the transcription of its gene is initiated from the normal initiation site, i.e. at the C at position −312, or from the heat-shock responding sites (Fig. 4).

The sequence corresponding to the presumable nuclear-targeting motif, RLIPR, found by Morishima et al. (5) in Ssc1

is not present in Ssp1 and the corresponding stretch is not conserved as such: instead, it is predicted to be RLINR from the nucleotide sequence data (22 and our unpublished results). Interestingly, the codons for the two R's in *S. cerevisiae* are both AGA, while those in *S. pombe* are both CGT. They are thus distinctly different, because these two codons belong to different families separated from each other a long time ago in their evolution. Nonetheless, the surrounding sequences are well conserved in the two yeasts. Thus, it could be that the nuclear-targeting signal differs in *S. cerevisiae* and *S. pombe* and that the sequence, RLINR, functions as one such signal in *S. pombe*.

## ACKNOWLEDGEMENTS

We are grateful to the Oriental Yeast Company (Osaka) for providing us with 10 kg of frozen yeast cells and to M.Yanagida (Kyoto University) for a *S.pombe* strain. This work was supported in part by grants-in-aid from the Ministry of Education, Science and Culture, Nos 62480465, 01880029, and 02954212.

## REFERENCES

1. Biswas, S.B. and Biswas, E.E. (1987) *J. Biol. Chem.*, **262**, 7831–7838.
2. Engman, D.M., Kirchhoff, L.V. and Donelson, J.E. (1989) *Mol. Cell. Biol.*, **9**, 5163–5168.
3. Welch, W.J. and Suhan, J.P. (1986) *J. Cell Biol.*, **103**, 2035–2052.
4. Haas, I.G. and Wabl, M. (1983) *Nature*, **306**, 387–389.
5. Morishima, N., Nakagawa, K., Yamamoto, E. and Shibata, T. (1990) *J. Biol. Chem.*, **265**, 15189–15197.
6. Lindquist, S. and Craig, E.A. (1988) *Ann. Rev. Genet.*, **22**, 631–677.
7. Craig, E.A., Kramer, J., Shilling, J., Werner-Washburne, M., Holmes, S., Kosic-Smithers, J. and Nicolet, C.M. (1989) *Mol. Cell. Biol.*, **9**, 3000–3008.
8. Kang, P.-J., Ostermann, J., Shilling, J., Neupert, W., Craig, E.A. and Pfanner, N. (1990) *Nature*, **348**, 137–143.
9. Hall, M.N., Hereford, L. and Herskowitz, I. (1984) *Cell*, **36**, 1057–1065.
10. Astell, C.R., Ahlstrom-Jonasson, L., Smith, M., Tatchell, K., Nasmyth, K.A. and Hall, B.D. (1981) *Cell*, **27**, 15–23.
11. Choe, J., Kolodrubetz, D. and Grunstein, M. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 1484–1487.
12. Wallis, J.W., Hereford, L. and Grunstein, M. (1980) *Cell*, **22**, 799–805.
13. Smith, M.M. and Andresson, O.S. (1983) *J. Mol. Biol.*, **169**, 663–690.
14. Matsushita, Y., Kitakawa, M. and Isono, K. (1989) *Mol. Gen. Genet.*, **219**, 119–124.
15. Kitakawa, M., Isono, K., Grohmann, L. and Graack, H.-R. (1990) *Nucl. Acids Res.*, **18**, 1521–1529.
16. Kang, W.-K., Matsushita, Y. and Isono, K. (1991) *Mol. Gen. Genet.*, **225**, 474–482.
17. Kang, W.-K., Matsushita, Y., Grohmann, L., Graack, H.-R., Kitakawa, M. and Isono, K. (1991) *J. Bacteriol.* **173**, 4013–4020.
18. Smith, D.E. and Fisher, P.A. (1984) *J. Cell Biol.*, **99**, 20–28.
19. Niederacher, D. and Entian, K.-D. (1987) *Mol. GeGenet.*, **206**, 505–509.
20. Sprague, G.F.Jr., Jensen, R. and Herskowitz, I. (1983) *Cell*, **32**, 409–415.
21. Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R. (1984) *Nucleic Acids Res.*, **12**, 7035–7056.
22. Powell, M.J. and Watts, F.Z. (1990) *Gene*, **95**, 105–110.
23. Boorstein, W.R. and Craig, E.A. (1990) *EMBO J.*, **9**, 2543–2553.
24. Scherer, P.E., Krieg, U.C., T.Hwang, S., Vestweber, D. and Schatz, G. (1990) *EMBO J.*, **9**, 4315–4322.
25. Zuker, M. and Stiegler, P. (1981) *Nucleic Acids Res.*, **9**, 133–148.
26. Pelham, H.R.B. (1982) *Cell*, **30**, 517–528.
27. Perisic, O., Xiao, H. and Lis, J.T. (1989) *Cell*, **59**, 797–806.
28. Parker, C.S. and Topol, J. (1984) *Cell*, **37**, 273–283.
29. Sorger, P.K. and Pelham, H.R.B. (1987) *EMBO J.*, **6**, 3035–3041.
30. Wiederrecht, G., Shuey, D.J., Kibbe, W.A. and Parker, C.S. (1987) *Cell*, **48**, 507–515.
31. Kingston, R.E., Schuetz, T.J. and Larin, Z. (1987) *Mol. Cell. Biol.*, **7**, 1530–1534.
32. Sorger, P.K., Lewis, M.J. and Pelham, H.R.B. (1987) *Nature*, **329**, 81–84.
33. Zimarino, V. and Wu, C. (1987) *Nature*, **327**, 727–730.
34. Jacobsen, B.K. and Pelham, H.R.B. (1988) *Mol. Cell. Biol.*, **8**, 5040–5042.
35. Nieto-Sotelo, J., Wiederrecht, G., Okuda, A. and Parker, C.S. (1990) *Cell*, **62**, 807–817.
36. Sorger, P.K. (1990) *Cell*, **62**, 793–805.
37. Kozak, M. (1987) *Nucleic Acids Res.*, **15**, 8125–8148.
38. Cigan, A.M. and Donahue, T.F. (1987) *Gene*, **59**, 1–18.
39. Kozak, M. (1988) *Mol. Cell. Biol.*, **8**, 2737–2744.
40. Pelham, H. (1985) *Trends Genet.*, **1**, 31–35.
41. McIntosh, E.M., Atkinson, T., Storms, R.K. and Smith, M. (1991) *Mol. Cell. Biol.*, **11**, 329–337.
42. McConnell, S.J., Stewart, L.C., Talin, A. and Yaffe, M.P. (1990) *J. Cell Biol.*, **111**, 967–976.
43. Gallo, G.J., Schuetz, T.J. and Kingston, R.E. (1991) *Mol. Cell. Biol.*, **11**, 281–288.
44. Kozak, M. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 8301–8305.
45. Natsoulis, G., Hilger, F. and Fink, G.R. (1986) *Cell*, **46**, 235–243.
46. Ellis, S.R., Morales, M.J., Li, J.-M., Hopper, A.K. and Martin, N.C. (1986) *J. Biol. Chem.*, **261**, 9703–9709.
47. Beltzer, J.P., Morris, S.R. and Kohlhaw, G.B. (1988) *J. Biol. Chem.*, **263**, 368–374.